Adaptive Filtering Algorithms for Quaternion-Valued Signals

by:
Sayedpouria Talebi

A thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy of Imperial College London

Communications and Signal Processing Research Group
Department of Electrical and Electronic Engineering
Imperial College London
2016

"The quaternion born, as a curios offspring of a quaternion of parents, say of geometry, algebra, metaphysics, and poetry ... I have never been able to give a clear statement of their nature and their aim than I have done in two lines of a sonnet addressed to Sir John Herschel:
"And how the one of Time, of Space the Three,
Might in the Chain of Symbols girdled be."
It is not so much to be wondered at, that they should have let me to strike out some new lines of research, which former methods have failed to suggest."

Sir William Rowan Hamilton

Abstract

Advances in sensor technology have made possible the recoding of three and four-dimensional signals which afford a better representation of our actual three-dimensional world than the "*flat view*" one and two-dimensional approaches. Although it is straightforward to model such signals as real-valued vectors, many applications require unambiguous modeling of orientation and rotation, where the division algebra of quaternions provides crucial advantages over real-valued vector approaches.

The focus of this thesis is on the use of recent advances in quaternion-valued signal processing, such as the quaternion augmented statistics, widely-linear modeling, and the $\mathbb{HR}$-calculus, in order to develop practical adaptive signal processing algorithms in the quaternion domain which deal with the notion of phase and frequency in a compact and physically meaningful way. To this end, first a real-time tracker of quaternion impropriety is developed, which allows for choosing between strictly linear and widely-linear quaternion-valued signal processing algorithms in real-time, in order to reduce computational complexity where appropriate. This is followed by the strictly linear and widely-linear quaternion least mean phase algorithms that are developed for phase-only estimation in the quaternion domain, which is accompanied by both quantitative performance assessment and physical interpretation of operations. Next, the practical application of state space modeling of three-phase power signals in smart grid management and control systems is considered, and a robust complex-valued state space model for frequency estimation in three-phase systems is presented. Its advantages over other available estimators are demonstrated both in an analytical sense and through simulations. The concept is then expanded to the quaternion setting in order to make possible the simultaneous estimation of the system frequency and its voltage phasors. Furthermore, a distributed quaternion Kalman filtering algorithm is developed for frequency estimation over power distribution networks and collaborative target tracking. Finally, statistics of stable quaternion-valued random variables, that include quaternion-valued Gaussian random variables as a special case, is investigated in order to develop a framework for the modeling and processing of heavy-tailed quaternion-valued signals.

Acknowledgment

I would like to thank my supervisor Professor Danilo P. Mandic for his supervision and technical support during my time at Imperial College London. I would also like to thank my family for all their unconditional support, especially my father Professor Sayed Morteza Talebi for his advice on matters of life and financial support without which I would not have been inspired or able to complete my education at this level.

Contents

Appendices:

List of Figures

List of Tables

Copyright

Statement of Originality

A significant part of this thesis is consisted of original contributions to the design and implementation of quaternion-valued adaptive filtering algorithms. As far as the author is aware, the work of other people has been acknowledge and referenced. The original contributions of this thesis are as follows:

- A real-time quaternion impropriety tracker with performance analysis in terms of its steady-state mean square error and bias.
- A quaternion least mean phase algorithm for adaptive phase-only estimation of quaternion-valued signals with performance analysis and a geometric interpretation of its operation.
- A widely-linear complex-valued frequency estimator for three-phase power distribution networks accounting for improper noise and the presence of harmonics.
- A quaternion frequency estimator for adaptive joint estimation of three-phase power system frequency and phasor information accounting for the presence of harmonics.
- A distributed quaternion Kalman filtering algorithm with applications to smart grids and collaborative target tracking.
- Investigation of quaternion-valued stable random variables and adaptive filtering of non-Gaussian processes using the augmented quaternion particle filter.

List of Publications

- S. P. Talebi, D. Xu, A. Kuh, and D. P. Mandic, "A Quaternion Least Mean Phase Adaptive Estimator," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 6419-6423, May 2014.

- S. P. Talebi and D. P. Mandic, "A Quaternion Frequency Estimator for Three-Phase Power Systems," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 3956-3960, April 2015.

- S. P. Talebi, S. Kanna, and D. P. Mandic, "Real-Time Estimation of Quaternion Impropriety," *In Proceedings of IEEE International Conference on Digital Signal Processing (DSP)*, pp. 557-561, July 2015.

- S. P. Talebi, S. Kanna, and D. P. Mandic, "A Non-Linear State Space Frequency Estimator for Three-Phase Power Systems," *In Proceeding of The International Joint Conference on Neural Networks (IJCNN)*, pp. 1-7, July 2015.

- S. P. Talebi, S. Kanna, and D. P. Mandic, "A Distributed Quaternion Kalman Filter with Applications to Smart Grid and Target Tracking," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 477-488, December 2016.

- S. P. Talebi, S. Kanna, Y. Xia, and D. P. Mandic, "A Distributed Quaternion Kalman Filter with Applications to Fly-by-Wire Systems," *To Appear In The Proceedings of IEEE International Conference on Digital Signal Processing (DSP)*, October 2016.

Abbreviations

ACLMS:	Augmented Complex Least Mean Square
AECKF:	Augmented Extended Complex Kalman Filter
AQCF:	Augmented Quaternion Characteristic Function
AQKF:	Augmented Quaternion Kalman Filter
AQPF:	Augmented Quaternion Particle Filter
AWGN:	Additive White Gaussian Noise
CAQKF:	Centralized Augmented Quaternion Kalman Filter
cdf:	cumulative distribution function
CLFE:	Complex Linear Frequency Estimator
CLMS:	Complex Least Mean Square
CNLFE:	Complex Non-Linear Frequency Estimator
CNLFE-WHC:	Complex Non-Linear Frequency Estimator With Harmonic Contamination
CWLFE:	Complex Widely-Linear Frequency Estimator
DAQKF:	Distributed Augmented Quaternion Kalman Filter
GNGD:	Generalized Normalized Gradient Descent
i.i.d:	independent and identically distributed
LMMP:	Least Mean Magnitude Phase
LMP:	Least Mean Phase
LMS:	Least Mean Square
LPF:	Low Pass Filter
MMSE:	Minimum Mean Square Error
MQFE:	Modified Quaternion Frequency Estimator
MSE:	Mean Square Error
NLMS:	Normalized Least Mean Square
pdf:	probability distribution function
PGNCS:	Primary Guidance, Navigation, and Control System

p.u.: per-unit

QAPE: Quaternion Adaptive Phasor Estimator

QFE: Quaternion Frequency Estimator

QHE: Quaternion Harmonic Estimator

QLMP: Quaternion Least Mean Phase

QLMS: Quaternion Least Mean Square

RLS: Recursive Least Squares

RMS: Root Mean Square

SNR: Signal to Noise Ratio

SSG: Sinusoidal Signal Generator

SαS: Symmetric α-Stable

WL-LMP: Widely-Linear Least Mean Phase

WL-QLMP: Widely-Linear Quaternion Least Mean Phase

WL-QLMS: Widely-Linear Quaternion Least Mean Square

Mathematical Notations

$\mathbb{R}$	Field of real numbers
$\mathbb{R}^+$	Field of positive real numbers
$\mathbb{C}$	Field of complex numbers
$\mathbb{H}$	Field of quaternion numbers
i, j, k	Imaginary units of the quaternion field
x	Scalers are denoted by lowercase letters
$\mathbf{x}$	Column vectors are denoted by bold lowercase letters
$\mathbf{X}$	Matrices are denoted by bold uppercase letters
X	Uni-variate random processes are denoted by uppercase Italic letters
$\boldsymbol{X}$	Multivariate random processes are denoted by uppercase bold Italic letters
$\mathbf{I}$	Identity matrix
μ	Adaptation gain
$(\cdot)^T$	Transpose operator
$(\cdot)^H$	Hermitian transpose operator
$(\cdot)^{-1}$	Matrix inverse operator
$(\cdot)^*$	Quaternion/complex conjugate operator
$\vert \cdot \vert$	Absolute value
$\Vert \cdot \Vert_p$	p^{th} norm
$\mathrm{Tr}(\cdot)$	Trace operator
$\det(\cdot)$	Determinant operator
$\mathrm{vec}(\cdot)$	Constructs a vector from a matrix by staking its columns
$(\cdot)^i$	i-involution operator
$(\cdot)^j$	j-involution operator
$(\cdot)^k$	k-involution operator
e	The Euler number
$\ln(\cdot)$	Natural logarithm

$\langle \cdot, \cdot \rangle$	The inner-product
$\times$	The cross-product
$E[\cdot]$	Statistical expectation operator
$E[y\vert x]$	Conditional expectation of y given x
$\mathbf{R}_{\mathbf{xy}}$	The cross-covariance operator, $\mathbf{R}_{\mathbf{xy}} = E[\mathbf{x}\mathbf{y}^T]$
$\mathbf{C}_{\mathbf{x}}$	The covariance operator, $\mathbf{C}_{\mathbf{x}} = E[\mathbf{x}\mathbf{x}^H]$
ρ_i, ρ_j, ρ_k	The i, j, and k quaternion impropriety measures
$\Re\{\cdot\}$	Real component of a quaternion/complex number
$\Im\{\cdot\}$	Imaginary component of a quaternion/complex number
$\Im_i\{\cdot\}$	i-imaginary component of a quaternion number
$\Im_j\{\cdot\}$	j-imaginary component of a quaternion number
$\Im_k\{\cdot\}$	k-imaginary component of a quaternion number
$\forall$	For all
$\triangleq$	Equal by definition
$\approx$	Approximately equal
$\propto$	Proportional to
∇	Gradient operator
∂	Partial differential operator
d	Total differential operator
$\hat{\mathbf{x}}^a_{n\vert n-1}$	The *a priori* estimate of $\mathbf{x}^a_n$
$\hat{\mathbf{x}}^a_{n\vert n}$	The *a posteriori* estimate of $\mathbf{x}^a_n$
n	Discrete time index
N	Number of signal dimensions
ΔT	Sampling interval
f_s	Sampling frequency
f	System frequency
$\hat{f}$	Estimate of system frequency
φ_n	Phase incrementing element
$P_{\boldsymbol{X}}(\cdot)$	The probability distribution function of $\boldsymbol{X}$
$F_{\boldsymbol{X}}(\cdot)$	The cumulative distribution function of $\boldsymbol{X}$
$\mathcal{U}[x, y)$	The uniform distribution on $[x, y)$

Chapter 1

Introduction

1.1 Overview

In contrast to digital filters with constant coefficients that are only optimal for those signals that they are specifically designed for, adaptive filters do not require any assumptions on the signal generating procedure and can operate optimally even in non-stationary environments [1, 2]. This has firmly put adaptive signal processing at the heart of statistical signal processing research since the 1950s with adaptive signal processing being used in a wide range of applications such as channel equalization, echo cancellation, magnetic resonance imaging, navigation instrumentation, radar, and sonar [1–3]. Moreover, in recent years an increasing amount of computational power has become readily available with lower purchasing and running costs than ever before, which has in turn paved the path for the deployment of adaptive filtering solutions for problems of growing complexity. In this chapter, a survey of those areas in adaptive signal processing that are relevant to the work in this thesis is presented. This helps clarify the motivations and contributions of the work in this thesis, which are summarized at the end of this chapter.

1.2 Signal processing in $\mathbb{R}$

The introduction of the recursive least squares (RLS) in the 1950s, mainly credited to Plackett [4], can be seen as the starting point of adaptive signal processing research in its modern form. The RLS algorithm is based on recursively finding the coefficients (weights) of a filter in a manner that minimizes a weighted summation of squares of an error measure. The use of this weighted summation of error measure squares however, makes for fast convergence rates on one hand and high computational complexity on the other. Although much research has been devoted to developing computationally efficient and fast versions of the RLS algorithm since its introduction [5–8], the computational complexity of the RLS algorithm remains as its major drawback when it comes to real-time signal processing applications. To this end, in 1959, Widrow and Hoff introduced their least mean square (LMS) algorithm [9, 10] while conducting research on adaptive pattern classification [9]. The LMS algorithm attempts to find optimal filtering coefficients by recursively updating the coefficients at each interval based on the gradient of the mean square error. A simple to implement and computationally efficient algorithm, the LMS has become the most well-known and widely used signal processing algorithm with many variants aimed at improving its performance including a class of linearly adjusted step-size LMS algorithms [11], the normalized least mean square (NLMS) algorithm that normalizes the input signal to guarantee convergence [12], and the generalized normalized gradient descent (GNGD) algorithm that adjusts the adaptation gain in a non-linear fashion [13, 14].

Approximately one year after the introduction of the LMS algorithm, Rudolf Kalman introduced his renowned Kalman filtering algorithm [15], based on the properties of conditional Gaussian random variables and linear processes which promptly found applications in the primary guidance, navigation, and control system (PGNCS) that guided the Apollo spacecrafts [16]. Since then, the framework has been expanded to cater for non-linearity in the form of the extended and unscented Klaman filters [17, 18], where the former linearizes the dynamic system equations around the current mean using the first-order Taylor series expansion whereas the later relies upon sigma-point linearization techniques. The Kalman filter and its non-linear variants have not only become the main stay of most navigation systems [16, 19–21], but also have found applications in areas such as finance, target tracking, robot vision, wireless communication, and neural networks [22–27].

For the Kalman filter to operate optimally, two conditions must be satisfied [28, 29]; the covariances of the observational and state evolution noise must be precisely established, which has resulted in specialized algorithms devoted to adaptive estimation of noise characteristics in the 1970s [29], and the underlying system model must be known, which has resulted in the introduction of dual Kalman filtering algorithms [30–32]. In essence dual Kalman filters implement two Kalman filtering algorithms with feedback, one employs the estimates of the system parameters in order to update the estimates of the process and the other employs estimates of the process to update the estimates of the system parameters.

Although signal processing in $\mathbb{R}$ has proven to be advantageous in a wide range of diverse applications extending from spacecraft navigation to finance, there are scenarios where it becomes necessary to use higher dimensional algebras. For instance, rotations in two and three-dimensional spaces can be presented in $\mathbb{R}^2$ and $\mathbb{R}^3$ using vector algebras; however, it becomes computationally expensive and analytically complicated to extract useful information from these matrices, such as the plane of motion, angle of rotation, and the speed of rotation (or rotation frequency). Furthermore, in a large number of applications, as will be discussed in Section 1.3 and Section 1.4, the complex and quaternion fields offer the dimensionality necessary to model the underlaying signal directly in the multi-dimensional domain that they naturally reside in, resulting in more elegant and efficient signal processing solutions.

1.3 Signal processing in $\mathbb{C}$

Invention is often times motivated by necessity and the invention of the complex field is no exception to this rule. Since the 8^{th} century mathematicians were interested in finding the roots of an arbitrary polynomial, most notably one can point to the work of Khwarizmi that provided a general solution for polynomials of up to the second degree with positive roots. However, the need for a number field beyond $\mathbb{R}$ became pressing in the 16^{th} century when a number of notable mathematicians, such as Niccolo Tartagila and Girolamo Cardano, where working on a closed form solution for polynomials of third and fourth order [33]. Rafael Bombelli, in his work on the roots of cubic polynomials introduced the symbol $\sqrt{-1}$ and showed that in order to solve the roots of such polynomials it is necessary to perform calculations in $\mathbb{C}$ [33]. Regarded as "*the most remarkable formula in mathematics*" by Richard Feynman, the utmost notable development regarding the use of complex numbers in engineering applications is the formula put forth by Euler, $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, that allows for a convenient presentation and a framework for analysis of two dimensional rotations, oscillation, and traveling waves [34–36].

In all fields of electrical engineering, complex-valued modeling of electrical components, circuits, and filters provides a simple mathematical framework for solving the differential equations that govern their behavior, allowing for speedy analysis of their performance [37, 38]. Therefore, most developments in these areas have taken place in the complex domain highlighting the need for complex-valued adaptive signal processing techniques. Hence, there has been a concerted effort to adapt real-valued algorithms for dealing with complex-valued signals. In 1975, Widrow presented the complex least mean square (CLMS) algorithm [39], which is a straightforward extension of the LMS algorithm to the complex field. The same approach has been taken in order to extend a number of real-valued signal processing techniques to the complex domain [3, 40–42], where often the covariance, $E[\mathbf{x}\mathbf{x}^T]$, is replaced with its complex dual, $E[\mathbf{x}\mathbf{x}^H]$, using the Hermitian operator instead of the transpose operator [43, 44]. Although a zero-mean real-valued Gaussian random variable is fully described by its covariance, this is not the case for zero-mean complex-valued Gaussian random variables as the covariance does not describe the inner-relation between the real and imaginary components of a complex-valued random variable [33, 43, 44]. In the case of complex-valued random variables, in order to fully describe the second-order information the covariance should be considered simultaneously with the pseudo-covariance, $E[\mathbf{x}\mathbf{x}^T]$ [33, 43, 44].

In order to better illustrate this fact, the heat map of the histogram of four different zero-mean complex-valued Gaussian random variables with unit variance is shown in Figure 1.1. Notice that although all four complex-valued Gaussian random variables are unit variance, the dependence structure of their real and imaginary components differs significantly.

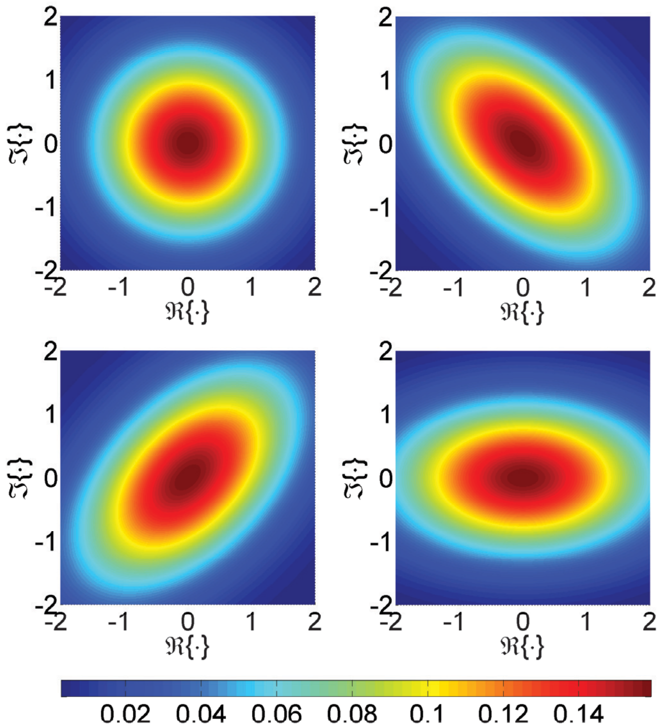

Figure 1.1: Heat map of the histogram of four zero-mean complex-valued Gaussian random variables with unit variance and differing pseudo-covariances.

In 2008, Javidi and Mandic introduced their augmented complex least mean square (ACLMS) algorithm [45, 46], that can exploit the full second-order statistical information of complex-valued signals. This was achieved by employing the $\mathbb{CR}$-calculus [47, 48] and the augmented complex statistics [49], where the complex-variable and its conjugate are treated as two different variables. The analysis indicates that the ACLMS outperforms the CLMS for complex-valued signals with a non-vanishing pseudo-covariance [33, 43, 44, 46]. The framework has also been deployed for optimal state space filtering of complex-valued signals by Goh and Mandic in order to develop an augmented extended complex Kalman filter (AECKF) [50], which since has been followed by a class of augmented Kalman filtering algorithms in [51, 52].

The Clarke transform [53] has been used for a long time in order to map the three-phase voltages onto the complex domain to simplify circuit analysis of three-phase power systems. More recently, augmented complex-valued signal processing techniques have been used for adaptive frequency estimation in three-phase power systems [54]. This is achieved by exploiting the Clarke transform to incorporate the information from all the phases and then employing the ACLMS or the AECKF to estimate the fundamental frequency of both balanced and unbalanced three-phase power systems [54–58]. Nonetheless, current algorithms can only model harmonic components as noise; in addition, their performance degrades when the three-phase systems is experiencing faults[1]. These issues are addressed in Chapter 5, where a new frequency estimator based on widely-linear complex-valued state space signal processing techniques is developed, its performance quantized, and compared to previous frequency estimation methods.

Although complex numbers have been instrumental in new developments in power, communications, and electrical engineering, mostly because of the platform they provide for solving differential equations, complex numbers lack the dimensionality necessary to model the actual three-dimensional world. For example, when dealing with three-phase power systems using the Clarke transform will lead to partial loss of information during faults. Therefore, it would be more elegant and analytically more appropriate to process these signals in the quaternion domain. This is further elaborated on in Section 1.4.

1.4 Signal processing in $\mathbb{H}$

First introduced by Hamilton [59], quaternions are a four-dimensional skew field and can be seen as the extension of complex numbers, where the imaginary component itself is three-dimensional. Since their introduction, quaternions have seen extensive use in physics, aerospace, and computer graphics [60–64] for describing orientation and rotation in three-dimensional spaces, due to their underlying division algebra and their natural ability to model three-dimensional data as pure imaginary quaternions [60–63]. In comparison to rotation matrices and vector algebras, quaternions present a more concise and computationally efficient representation of three-dimensional rotations that allows to avoid problems associated with gimbal lock [60], resulting in accurate and mathematically tractable solution with fewer constrains than those obtained by vector algebras in $\mathbb{R}^3$.

In signal processing applications, quaternions have only come to prominence since the 1990s finding applications in processing of colored images [65–67], owing to their four-dimensional nature, and in attitude estimation for aerospace control systems [62], owing to their ability to model three-dimensional rotations. However, due to the lack of a suitable mathematical framework, most signal processing algorithms developed for quaternion-valued signals have taken the approach of essentially transforming quaternions into a vector of quadrivariate real-valued components [62–64, 68], where most advantages of quaternions such as their division algebra and their underlying physical interpretation are lost. Suitable tools for developing rigorous signal processing algorithms, such as the $\mathbb{HR}$-calculus [69, 70], the augmented second-order statistics of

[1]A three-phase system that does not have balanced voltage phasors in its main frequency component is referred to as operating under fault conditions or experiencing a fault.

quaternions [71], and the quaternion fast Fourier transform [72, 73], have only recently been developed sparking a resurgence of quaternions in signal processing with signal processing algorithms that are derived directly in $\mathbb{H}$ proving advantageous in kernel learning [74], wind profile forecasting [75], and bearings-only tracking [76].

1.4.1 Augmented second-order statistics of quaternions

At around the same time Took *et al.* [71] and Via *et al.* [77, 78] revisited the second-order statistics of quaternion random variables establishing a mathematical framework for exploiting the full second-order information of quaternion-valued signals. This was achieved by considering the quaternion signal and its rotations around the three axes of the imaginary quaternion subspace simultaneously. Similar to the augmented complex statistics [49], it is shown that the standard covariance only partially explains the second-order statistics of quaternion-valued signals as it does not reveal any information about the dependence structure of the real-valued components of quaternion-valued signals.

The dependence structure of the real-valued components of a quaternion-valued signal is revealed by the cross-covariance between the quaternion-valued signal and its rotations around the three axes of the quaternion imaginary subspace, hereafter referred to as pseudo-covariances [71], which has in turn allowed for the establishment of the widely-linear model and its optimal "*Wiener type*" solution for conditional minimum mean square error (MMSE) estimation in the quaternion domain [71]. In addition, it has been shown that the widely-linear model can be simplified reducing computational complexity of adaptive signal processing algorithms if the quaternion-valued signal has two or three vanishing pseudo-covariances [77, 78].

1.4.2 The $\mathbb{HR}$-calculus

One major stumbling block when it comes to developing adaptive signal processing algorithms in the quaternion domain has been the restrictiveness of the Cauchy-Riemann-Fueter condition for quaternion differentiability, that can only accommodate for constants and linear functions [79]. Therefore, in 2010 and following along the same path as the $\mathbb{CR}$-calculus [47, 48], the $\mathbb{HR}$-calculus was developed by Janhanchahi, Took, and Mandic [69, 70, 79]. The $\mathbb{HR}$-calculus establishes a duality between $\mathbb{H}^4$ and $\mathbb{R}^4$ by rotating the imaginary part of the quaternion-valued variable around the axes of the quaternion imaginary subspace, much like the augmented quaternion statistics, resulting in a framework for calculating the derivatives of quaternion-valued functions directly in the quaternion domain [69, 70, 79]. The sections of the $\mathbb{HR}$-calculus that are relevant to this thesis are explained in detail in Chapter 2.

1.5 Motivation and contributions

Although a great deal of research has been carried out on signal processing in $\mathbb{H}$, practical quaternion-valued signal processing algorithms dealing with the notion of phase and frequency in the quaternion domain are still lacking. Motivated by the recent developments in quaternion-valued signal processing and the natural ability of quaternions to model three-dimensional rotations, we focus on quaternion-valued signal processing algorithms dealing with the concept of phase and frequency in the quaternion domain. The contributions of this thesis are summarized as follows:

- A novel real-time tracker of quaternion impropriety is developed. This allows to track the dependence structure between the real-valued components of quaternion-valued random variables in real-time, that in turn provides useful information on selecting between widely-linear, semi-widely-linear, or strictly linear versions of adaptive signal processing algorithms and reducing computational complexity.
- A quaternion-valued phase-only estimator is developed for adaptive estimation of the phase of quaternion-valued signals. The work includes both a qualitative performance assessment and a physical interpretation of the operations of the developed algorithm.
- Although a number of strictly linear and widely-linear complex-value signal processing algorithms for adaptive frequency estimation in three-phase power systems have already been introduced; However, the performance of these algorithms degrades rapidly in crucial moments when operating under fault conditions. To this end, a widely-linear complex-valued frequency estimator for three-phase systems that can outperform its counterparts in addition to having consistent performance under both nominal and fault conditions is developed.
- It is shown that mapping the three-phase signal onto the complex domain leads to partial loss of information when operating under fault conditions; therefore, a quaternion frequency estimator for three-phase systems that can accommodate for the three-phase signal without loss of information is developed. The developed frequency estimator not only outperforms its complex-valued counterparts, but also provides a platform for adaptive estimation of the system voltage phasors. Finally, the framework is expanded to account for the presence of harmonics in the power system.
- The work on quaternion state space estimators is expanded to the distributed setting with the development of the distributed quaternion Kalman filter. The performance analysis of the developed algorithm indicates that it operates in an unbiased fashion. In addition, the mean square error performance of the developed algorithm is quantified.

- Most of the research surrounding quaternion-valued signal processing is concentrated around Gaussian random variables, due to their stable property. In order to develop a framework for modeling and processing of heavy-tailed quaternion-valued signals, the statistical analysis of quaternion-valued random variables through the characteristic function is considered, which allows a particle filtering algorithm for adaptive processing of elliptically contoured stable quaternion-valued random variables, that include quaternion-valued Gaussian random variables as a special case, to be proposed.

1.6 Organization of thesis

In order to better clarify our contributions, an overview of those areas in adaptive signal processing that are relevant to this thesis was presented in Chapter 1. In Chapter 2, the theoretical background which forms the mathematical foundations of this thesis, including the $\mathbb{HR}$-calculus and the augmented second-order statistics of quaternion-valued random variables, are presented so that the reader can better understand the work conducted in this thesis and included in the following chapters.

In Chapter 3, a novel algorithm for tracking quaternion impropriety in real-time is developed and its performance under different degrees of impropriety is analyzed. In Chapter 4, adaptive phase estimation of quaternion-valued signals is considered, where a quaternion phase only estimator is developed, its performance analyzed, and its operation explained from a geometric point of view. Frequency estimation in three-phase power systems is looked into in Chapter 5, where a complex-valued adaptive frequency estimator that can outperform its complex-valued counterparts, account for presence of harmonics, and has consistent performance under nominal and fault conditions is developed. Frequency estimation in three-phase power systems is considered in the quaternion domain in Chapter 6, in order to incorporate all the available information in the voltage recordings and estimate the system voltage phasors. In Chapter 7, a distributed quaternion Kalman filtering algorithm with applications to smart grid and collaborative target tracking is developed and its mean and mean square error performance analyzed. In Chapter 8, the generality of quaternion-value stable random variables are considered in order to establish a framework for modeling and processing of heavy-tailed quaternion-valued signals. Finally, the work is concluded in Chapter 9, where the contributions of this thesis and recommendations for potential future work are summarized.

Chapter 2

Background

2.1 Overview

In this chapter the mathematical foundations of signal processing in $\mathbb{C}$ and $\mathbb{H}$, such as the analyticity and differentiability of complex and quaternion-valued functions and the augmented statistics of complex and quaternion-valued random variables, are revised.

2.2 Analyticity and differentiability in $\mathbb{C}$

Consider a complex-valued function $f(z) = f_r(z_r, z_i) + if_i(z_r, z_i)$, where $z \in \mathbb{C}$ with f_r and f_i denoting the real and imaginary parts of f, whereas the real and imaginary parts of z are denoted by z_r and z_i. The derivative of f is given by

$$\frac{\partial f}{\partial z} = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} \tag{2.1}$$

where if for a given point on the complex plane, $z = z_0$, the limit in (2.1) exists and is independent of the direction taken, then f is said to be differentiable at $z = z_0$. In addition, in the complex domain, the function f is differentiable at $z = z_0$ if and only if the partial differentials of f_r and f_i exist and satisfy the Cauchy-Riemann conditions given by [80]

$$\frac{\partial f_r}{\partial z_r} = \frac{\partial f_i}{\partial z_i} \qquad \text{and} \qquad \frac{\partial f_i}{\partial z_r} = -\frac{\partial f_r}{\partial z_i}.$$

Moreover, if f admits the Cauchy-Riemann conditions in some neighborhood of $z = z_0$, then f is infinitely differentiable at $z = z_0$, can be locally expanded in a power series, and is referred to as analytic [80]. However, the functions encountered in signal processing applications are often not analytic. One such example is $f = zz^* = z_r^2 + z_i^2$, which is used as the cost function of the CLMS and ACLMS algorithms [45, 46] and is not analytic at any point on the complex plane, that is with the exception of $z = 0$. Hence, the Cauchy-Riemann conditions constitute a very strict structure on differentiable complex-valued functions.

In the context of the $\mathbb{CR}$-calculus [81], $f : \mathbb{C} \to \mathbb{C}$ is considered as a bivariate function $h(z_r, z_i) = [f_r, if_i]^T$, where the total differential of h is given by

$$\begin{aligned} \mathrm{d}h =& \frac{\partial h}{\partial z_r}\mathrm{d}z_r + \frac{\partial h}{\partial z_i}\mathrm{d}z_i \\ =& \frac{\partial f_r}{\partial z_r}\mathrm{d}z_r + i\frac{\partial f_i}{\partial z_r}\mathrm{d}z_r + \frac{\partial f_r}{\partial z_i}\mathrm{d}z_i + i\frac{\partial f_i}{\partial z_i}\mathrm{d}z_i \end{aligned} \tag{2.2}$$

where, after some tedious mathematical manipulations and applying the transforms

$$\mathrm{d}z_r = (\mathrm{d}z + \mathrm{d}z^*)/2 \qquad \text{and} \qquad \mathrm{d}z_i = (\mathrm{d}z - \mathrm{d}z^*)/2i$$

the expression in (2.2) can be simplified to give [81]

$$\mathrm{d}h = \frac{1}{2}\left(\frac{\partial h}{\partial z_r} - i\frac{\partial h}{\partial z_i}\right)\mathrm{d}z + \frac{1}{2}\left(\frac{\partial h}{\partial z_r} + i\frac{\partial h}{\partial z_i}\right)\mathrm{d}z^*. \tag{2.3}$$

Now, by slightly abusing the mathematics, considering z and z^* as "*algebraically*" independent[1] variables, results in $f(z_r, z_i) \triangleq f(z, z^*)$, where the total differential of f is given by

$$\mathrm{d}f = \frac{\partial f}{\partial z}\mathrm{d}z + \frac{\partial f}{\partial z^*}\mathrm{d}z^*. \tag{2.4}$$

It follows from comparing the expressions in (2.3) and (2.4) that

$$\begin{aligned} \frac{\partial f}{\partial z} &= \frac{1}{2}\left(\frac{\partial f}{\partial z_r} - i\frac{\partial f}{\partial z_i}\right) \\ \frac{\partial f}{\partial z^*} &= \frac{1}{2}\left(\frac{\partial f}{\partial z_r} + i\frac{\partial f}{\partial z_i}\right) \end{aligned}$$

where $\partial f/\partial z$ is referred to as the $\mathbb{R}$-derivative whereas $\partial f/\partial z^*$ is referred to as the $\mathbb{R}^*$-derivative [81]. In addition, since z and z^* were considered as "*algebraically*" independent, the partial derivatives $\partial f/\partial z$ and $\partial f/\partial z^*$ can be calculated in the same fashion as for real-valued multi-variate functions with z and z^* being considered as different variables, resulting in $\partial z/\partial z^* = 0$ and $\partial z^*/\partial z = 0$, which is the reason behind referring to z and z^* as "*algebraically*" independent. Note that if the Cauchy-Riemann conditions are satisfied, then $\partial f/\partial z^* = 0$ and the total derivative of f will be equal to its $\mathbb{C}$-derivative, $\partial f/\partial z$ [81].

[1] Although z and z^* are essentially the same variable, treating them as different variables provides a platform where the two degrees of freedom inherent in complex numbers can be accommodated for.

2.3 Augmented estimation in $\mathbb{C}$

It has now become generally accepted that for a complex-valued random variable, $\mathbf{x}$, the standard covariance, $E[\mathbf{x}\mathbf{x}^H]$, does not fully describe its second-order statistical information [33, 82]. The full description of the second-order statistical information of a general complex-valued random variable is only possible through the augmented complex statistics, where the complex random variable, $\mathbf{x}$, is augmented with its conjugate, $\mathbf{x}^*$ to give the augmented random variable as $\mathbf{x}^a = [\mathbf{x}^T, \mathbf{x}^H]^T$ [33]. Now, the augmented covariance matrix can be expressed as

$$\mathbf{C}_{\mathbf{x}^a} = E[\mathbf{x}^a \mathbf{x}^{aH}] = \begin{bmatrix} \mathbf{C}_{\mathbf{x}} & \mathbf{R}_{\mathbf{xx}} \\ \mathbf{R}_{\mathbf{xx}}^H & \mathbf{C}_{\mathbf{x}}^H \end{bmatrix}$$

where $E[\cdot]$ represents the statistical expectation, while $\mathbf{C}_{\mathbf{x}} = E[\mathbf{x}\mathbf{x}^H]$ is the standard covariance and $\mathbf{R}_{\mathbf{xx}} = E[\mathbf{x}\mathbf{x}^T]$ is referred to as the pseudo-covariance [33]. It now becomes apparent that the standard covariance can fully describe the second-order statistical information of only a spacial class of complex-valued random-variables that have vanishing pseudo-covariances, for which the probability distribution is rotation invariant and are referred to as second-order proper[2]; however, for a general complex-valued random variable both the covariance and the pseudo-covariance are required to fully exploit their second-order statistics [33].

In order to introduce the optimal second-order estimator for complex-valued signals, first we must revisit the real-valued MMSE estimator that estimates y conditional to observation x, given by

$$\hat{y} = E[y|x]$$

where $\hat{y}$ is the estimate of y. For zero-mean and jointly Gaussian x and y the optimal solution is the strictly linear estimator given by

$$\hat{y} = \mathbf{h}^T \mathbf{x}$$

where $\mathbf{h}$ is a vector of coefficients and $\mathbf{x}$ is a vector of past observations referred to as the regressor. Now, in the case where x and y are complex-valued, the MMSE estimator should be expressed in terms of the real and imaginary components of x and y [33], which yields

$$\hat{y} = \underbrace{E[y_r|x_r, x_i]}_{\hat{y}_r} + i\,\underbrace{E[y_i|x_r, x_i]}_{\hat{y}_i} \tag{2.5}$$

[2]In the literature the phrases "*second-order proper*" and "*second-order circular*" are often used interchangeably [33, 43, 44, 83]; However, in this thesis we will refer to complex or quaternion-valued random variables with vanishing pseudo-covariances as proper. In addition, the words "*second-order*" are often dropped as Gaussianity is assumed.

where upon replacing $x_r = (x + x^*)/2$ and $x_i = i(x^* - x)/2$ we have

$$\hat{y} = E[y_r|x, x^*] + iE[y_i|x, x^*].$$

Therefore, the optimal MMSE estimator for complex-valued zero-mean and jointly Gaussian x and y becomes

$$\hat{y} = \mathbf{h}^H\mathbf{x} + \mathbf{g}^H\mathbf{x}^* \tag{2.6}$$

where $\mathbf{h}$ and $\mathbf{g}$ are complex-valued coefficient vectors. From Section 2.2, recall that $\partial x/\partial x^* = \partial x^*/\partial x = 0$, which results in the estimator in (2.6) being referred to as widely-linear, due to the fact that it is linear with respect to both $\mathbf{x}$ and $\mathbf{x}^*$. In addition, the estimator in (2.6) can be rearranged into a more elegant representation as

$$\underbrace{\begin{bmatrix} \hat{y} \\ \hat{y}^* \end{bmatrix}}_{\hat{\mathbf{y}}^a} = \underbrace{\begin{bmatrix} \mathbf{h}^H & \mathbf{g}^H \\ \mathbf{g}^T & \mathbf{h}^T \end{bmatrix}}_{\mathbf{W}^a} \underbrace{\begin{bmatrix} \mathbf{x} \\ \mathbf{x}^* \end{bmatrix}}_{\mathbf{x}^a}$$

where $\hat{\mathbf{y}}^a$ and $\mathbf{x}^a$ are the augmented estimation and augmented regressor vectors, while $\mathbf{W}^a$ is the augmented weight matrix. Since the widely-linear estimator uses the regressor vector in its augmented form, it is equivalently referred as the augmented estimator.

The augmented statistics of complex-valued signals and the $\mathbb{CR}$-calculus have been exploited in [51] to introduce a class of augmented complex Kalman filtering algorithms including the AECKF. To better explain the operations of the AECKF, consider the evolution sequence of the complex-valued augmented state vector $\{\mathbf{x}_n^a, n = 0, 1, 2, \ldots\}$, given by

$$\mathbf{x}_n^a = f_n(\mathbf{x}_{n-1}^a) + \boldsymbol{\nu}_n^a$$

where $f_n(\cdot)$ is the state evolution function at time instant n and $\{\boldsymbol{\nu}_n^a, n = 0, 1, 2, \ldots\}$ is the augmented state evolution noise sequence. In Kalman filtering, the objective is to track $\mathbf{x}_n^a$, in real-time, through noise corrupted observations that are expressed as

$$\mathbf{y}_n^a = h_n(\mathbf{x}_n^a) + \boldsymbol{\omega}_n^a$$

where $\mathbf{y}_n^a$ and $h_n(\cdot)$ are the augmented observation vector and observation function at time instant n, while $\{\boldsymbol{\omega}_n^a, n = 0, 1, 2, \cdots\}$ is the augmented measurement noise sequence. Hereafter, at each time instant, the observation and state evolution functions are approximated as $f_n(\mathbf{x}_n^a) \approx \mathbf{A}_n^a\mathbf{x}_n^a$ and $h_n(\mathbf{x}_n^a) \approx \mathbf{H}_n^a\mathbf{x}_n^a$, where $\mathbf{A}_n^a$ and $\mathbf{H}_n^a$ are the Jacobian matrices of $\mathrm{f}_n(\cdot)$ and $h_n(\cdot)$, obtained through the framework of the $\mathbb{CR}$-calculus. The augmented state vector sequence can now be tracked using the AECKF given in Algorithm 1, where $\mathbf{C}_{\boldsymbol{\nu}_n^a}$ and $\mathbf{C}_{\boldsymbol{\omega}_n^a}$ denote the augmented covariance matrices of $\boldsymbol{\nu}_n^a$ and $\boldsymbol{\omega}_n^a$, while $\hat{\mathbf{x}}_{n|n-1}^a$ and $\hat{\mathbf{x}}_{n|n}^a$ represent the *a priori* and *a posteriori* estimates of $\mathbf{x}_n^a$, whereas $\hat{\mathbf{M}}_{n|n-1}^a$ and $\hat{\mathbf{M}}_{n|n}^a$ are estimates of the augmented covariance matrices of the *a priori* and *a posteriori* state estimation error vectors.

Algorithm 1. AECKF [51]

Initialize with:

$$\hat{\mathbf{x}}^a_{0|0} = E[\mathbf{x}^a_0]$$
$$\hat{\mathbf{M}}^a_{0|0} = E\big[(\mathbf{x}^a_0 - E[\mathbf{x}^a_0])(\mathbf{x}^a_0 - E[\mathbf{x}^a_0])^H\big]$$

Model update:

$$\hat{\mathbf{x}}^a_{n|n-1} = f_n(\hat{\mathbf{x}}^a_{n-1|n-1})$$
$$\hat{\mathbf{M}}^a_{n|n-1} = \mathbf{A}^a_n \hat{\mathbf{M}}^a_{n-1|n-1} \mathbf{A}^{aH}_n + \mathbf{C}_{\boldsymbol{\nu}^a_n}$$

Measurement update:

$$\mathbf{G}^a_n = \hat{\mathbf{M}}^a_{n|n-1} \mathbf{H}^{aH}_n \left(\mathbf{H}^a_n \hat{\mathbf{M}}^a_{n|n-1} \mathbf{H}^{aH}_n + \mathbf{C}_{\boldsymbol{\omega}^a_n}\right)^{-1}$$
$$\hat{\mathbf{x}}^a_{n|n} = \hat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_n \left(\mathbf{y}^a_n - h_n(\hat{\mathbf{x}}^a_{n|n-1})\right)$$
$$\hat{\mathbf{M}}^a_{n|n} = \big(\mathbf{I} - \mathbf{G}^a_n \mathbf{H}^a_n\big) \hat{\mathbf{M}}^a_{n|n-1}$$

2.4 Quaternion algebra

The skew-field of quaternions is a four-dimensional, non-commutative, associative, division algebra. A quaternion variable $q \in \mathbb{H}$ consists of a real part, $\Re(q)$, and a three-dimensional imaginary part or pure quaternion, $\Im(q)$, that is also referred to as the vector part due to the fact that it comprises three components $\Im_i(q)$, $\Im_j(q)$, and $\Im_k(q)$; hence, q can be expressed as

$$\begin{aligned} q = \Re(q) + \Im(q) =& \Re(q) + \Im_i(q) + \Im_j(q) + \Im_k(q) \\ =& q_r + iq_i + jq_j + kq_k \end{aligned}$$

where $q_r, q_i, q_j, q_k \in \mathbb{R}$. The unit vectors i, j, and k are the orthonormal basis for the quaternion imaginary subspace and obey the following product rules

$$\begin{gathered} ij = k, jk = i, ki = j, \\ i^2 = j^2 = k^2 = ijk = -1. \end{gathered}$$

The product of $q_1, q_2 \in \mathbb{H}$ is given by

$$\begin{aligned} q_1 q_2 =& \Re(q_1)\Re(q_2) + \Re(q_1)\Im(q_2) + \Re(q_2)\Im(q_1) + \\ & \Im(q_1) \times \Im(q_2) - \big\langle \Im(q_1), \Im(q_2) \big\rangle \end{aligned} \tag{2.7}$$

where the symbols '$\langle \cdot, \cdot \rangle$' and '$\times$' denote the inner and cross-products, respectively. Moreover, notice that due to the cross-product in (2.7), the quaternion product is non-commutative unless q_1 and q_2 have parallel imaginary components.

The involution of $q \in \mathbb{H}$ around $\zeta \in \mathbb{H}$ is defined as $q^{\zeta} \triangleq \zeta q \zeta^{-1}$ [84] and can be seen as the quaternion equivalent of the complex conjugate, as the real-valued components of a quaternion number, $q \in \mathbb{H}$, can be expressed using involutions as [69, 71, 79]

$$\begin{aligned} q_r =& \frac{1}{4}\left(q + q^i + q^j + q^k\right) \quad & q_i =& \frac{1}{4i}\left(q + q^i - q^j - q^k\right) \\ q_j =& \frac{1}{4j}\left(q - q^i + q^j - q^k\right) \quad & q_k =& \frac{1}{4k}\left(q - q^i - q^j + q^k\right). \end{aligned} \tag{2.8}$$

Furthermore, in the case that $\zeta^2 = -1$ the following expressions hold true

$$\begin{aligned} (q^{\zeta})^* =& (q^*)^{\zeta} \\ \zeta(q^{\zeta}) =& q\zeta \\ (q^{\zeta})^{\zeta} =& q. \end{aligned} \tag{2.9}$$

The quaternion conjugate is also an involution and is defined as

$$q^* = \Re(q) - \Im(q) = \frac{1}{2}\left(q^i + q^j + q^k - q\right) \tag{2.10}$$

while the norm of $q \in \mathbb{H}$ is given by

$$|q| = \sqrt{qq^*} = \sqrt{q_r^2 + q_i^2 + q_j^2 + q_k^2}$$

whereas the quaternion inverse can be written as $q^{-1} = q^*/|q|$. In addition, $\forall\{q_1, q_2\} \in \mathbb{H}$ the following properties hold

$$\begin{aligned} |q_1 q_2| =& |q_1||q_2| \\ \left|\frac{q_1}{q_2}\right| =& \frac{|q_1|}{|q_2|} \\ (q_1 q_2)^* =& q_2^* q_1^*. \end{aligned}$$

In a similar fashion to complex numbers, a quaternion $q \in \mathbb{H}$ can alternatively be expressed by its polar presentation, given by [72]

$$q = |q|e^{\xi\theta} = |q|\big(\cos(\theta) + \xi\sin(\theta)\big)$$

where

$$\xi = \frac{\Im(q)}{|\Im(q)|} \qquad \text{and} \qquad \theta = \operatorname{atan}\left(\frac{|\Im(q)|}{\Re(q)}\right).$$

Moreover, it is straightforward to prove that the $\sin(\cdot)$ and $\cos(\cdot)$ functions can be expressed as

$$\sin(\theta) = \frac{1}{2\xi}\left(e^{\xi\theta} - e^{-\xi\theta}\right) \quad \text{and} \quad \cos(\theta) = \frac{1}{2}\left(e^{\xi\theta} + e^{-\xi\theta}\right) \tag{2.11}$$

where $\xi^2 = -1$. Note that to express the $\sin(\cdot)$ and $\cos(\cdot)$ functions in their polar from, as in (2.11), ξ can be replaced with an arbitrary normalized pure quaternion number [72].

2.5 Three-dimensional rotations

The theorem of rigid-body rotations put forth by Euler states that the attitude of a body after having foregone any sequence of rotations is equivalent to a single right-hand rotation of that body by an angle θ about an axis $\boldsymbol{\eta}$ parallel to the direction that is unchanged by the rotation [60, 61]. Traditionally three-dimensional rotations are represented using rotation matrices, where a rotation of θ degrees around the unit vector $\boldsymbol{\eta} = [\eta_x, \eta_y, \eta_z]^T$ is expressed as $\mathbf{u}' = \mathbf{R}\mathbf{u}$ with the vectors $\mathbf{u}$ and $\mathbf{u}'$ representing the pre and post-rotation coordinates of the rigid-body, whereas the rotation matrix $\mathbf{R}$ is given by

$$\mathbf{R} = \begin{bmatrix} \cos(\theta) + \eta_x^2 (1 - \cos(\theta)) & \eta_x\eta_y (1 - \cos(\theta)) - \eta_z \sin(\theta) & \eta_x\eta_z (1 - \cos(\theta)) + \eta_y \sin(\theta) \\ \eta_y\eta_x (1 - \cos(\theta)) + \eta_z \sin(\theta) & \cos(\theta) + \eta_y^2 (1 - \cos(\theta)) & \eta_y\eta_z (1 - \cos(\theta)) - \eta_x \sin(\theta) \\ \eta_z\eta_x (1 - \cos(\theta)) - \eta_y \sin(\theta) & \eta_z\eta_y (1 - \cos(\theta)) + \eta_x \sin(\theta) & \cos(\theta) + \eta_z^2 (1 - \cos(\theta)) \end{bmatrix} \tag{2.12}$$

while the relation between the pre ans post-rotation coordinates can more conveniently be represented as

$$\mathbf{u}' = \cos(\theta)\mathbf{u} + \sin(\theta)\left(\boldsymbol{\eta} \times \mathbf{u}\right) + \left(1 - cos(\theta)\right)\eta\eta^T\mathbf{u} \tag{2.13}$$

with the expression '$\boldsymbol{\eta} \times \mathbf{u}$' denoting cross-product of the vectors $\boldsymbol{\eta}$ and $\mathbf{u}$ [61]. It now becomes apparent that in addition to giving a counterintuitive depiction of the rotation operation, extracting the parameters θ and $\boldsymbol{\eta}$ from $\mathbf{R}$ is a computationally expensive and an inconvenient affair.

Modeling the pre-rotation Cartesian coordinates, $\mathbf{u} = [u_x, u_y, u_z]^T$, as a pure quaternion given by $q_{\mathbf{u}} = iu_x + ju_y + ku_z$; then, the post-rotation coordinates can be calculated through the involution

$$q_{\mathbf{u}'} = \left(\cos\left(\frac{\theta}{2}\right) + q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\right) q_{\mathbf{u}} \left(\cos\left(\frac{\theta}{2}\right) + q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\right)^{-1}$$

where $q_{\mathbf{u}'} = iu'_x + ju'_y + ku'_z$ represents the post-rotation coordinates, while $q_{\boldsymbol{\eta}} = i\eta_x + j\eta_y + k\eta_z$, with the rotation fully characterized by $\big(\cos(\theta/2) + q_{\boldsymbol{\eta}}\sin(\theta/2)\big) = e^{q_{\boldsymbol{\eta}}\theta/2}$, much like the Euler formula that models two-dimensional rotations.

The advantages of modeling three-dimensional rotations employing quaternions as compared to rotation matrices are summarized in the following [60, 61, 85, 86]:

- The rotation matrix, $\mathbf{R}$, requires nine variables to express a rotation, whereas its quaternion equivalent, $e^{q_{\boldsymbol{\eta}}\theta/2}$, only requires four variables when modeling the same rotation, which reduces both memory requirements and memory access time by more than half.

- A rotation matrix must satisfy the conditions $\mathbf{R}\mathbf{R}^T = \mathbf{I}$ and $\det(\mathbf{R}) = 1$; therefore, the need might arise to calibrated the rotation matrix after many rotations have been performed, due to finite precision of computer calculations, which is computationally expensive, whereas no such operation is required when modeling rotations with quaternions [85].

- Expressing a sequence of rotations applying the roll, pitch, and yaw angles, a degree of freedom is lost when one of the angles reaches $\pi/2$. However, this is not the case for quaternions where only the angle and the axis of rotation are required.

- It is straightforward to produce smooth interpolations of rotations when they are modeled with quaternions allowing for higher quality computer graphics [86].

2.6 The $\mathbb{HR}$-calculus

The Cauchy-Riemann-Fueter condition for differentiability in $\mathbb{H}$, given by [79, 85]

$$\frac{\partial f}{\partial q_r} + i\frac{\partial f}{\partial q_i} + j\frac{\partial f}{\partial q_j} + k\frac{\partial f}{\partial q_k} = 0$$

pose a severe restriction on the class of quaternion-valued differentiable functions, as it only accommodates for constant or linear functions in $\mathbb{H}$. This has been a major stumbling block in the derivation of quaternion-valued signal processing algorithms, as in most adaptive signal processing techniques the aim is to minimize a cost function of an error measure which is typically a real-valued function of quaternion-valued variables.

One elegant solution to this problem is the $\mathbb{HR}$-calculus [69, 70, 79, 85]. In the context of the $\mathbb{HR}$-calculus, a quaternion function $f(q) : \mathbb{H} \to \mathbb{H}$ represented as

$$f(q) = f_r(q_r, q_i, q_j, q_k) + if_i(q_r, q_i, q_j, q_k) + jf_j(q_r, q_i, q_j, q_k) + kf_j(q_r, q_i, q_j, q_k)$$

in a format similar to that of the $\mathbb{CR}$-calculus, is expressed as a quadrivariate function $g = [f_r, if_i, jf_j, kf_k]^T$ with the total differential

$$\begin{aligned}
\mathrm{dg} =& \frac{\partial \mathrm{g}}{\partial q_r}\mathrm{d}q_r + \frac{\partial \mathrm{g}}{\partial q_i}\mathrm{d}q_i + \frac{\partial \mathrm{g}}{\partial q_j}\mathrm{d}q_j + \frac{\partial \mathrm{g}}{\partial q_k}\mathrm{d}q_k \\
=& \frac{\partial \mathrm{f}_r}{\partial q_r}\mathrm{d}q_r + i\frac{\partial \mathrm{f}_i}{\partial q_r}\mathrm{d}q_r + j\frac{\partial \mathrm{f}_j}{\partial q_r}\mathrm{d}q_r + k\frac{\partial \mathrm{f}_k}{\partial q_r}\mathrm{d}q_r \\
&+ \frac{\partial \mathrm{f}_r}{\partial q_i}\mathrm{d}q_i + i\frac{\partial \mathrm{f}_i}{\partial q_i}\mathrm{d}q_i + j\frac{\partial \mathrm{f}_j}{\partial q_i}\mathrm{d}q_i + k\frac{\partial \mathrm{f}_k}{\partial q_i}\mathrm{d}q_i \\
&+ \frac{\partial \mathrm{f}_r}{\partial q_j}\mathrm{d}q_j + i\frac{\partial \mathrm{f}_i}{\partial q_j}\mathrm{d}q_j + j\frac{\partial \mathrm{f}_j}{\partial q_j}\mathrm{d}q_j + k\frac{\partial \mathrm{f}_k}{\partial q_j}\mathrm{d}q_j \\
&+ \frac{\partial \mathrm{f}_r}{\partial q_k}\mathrm{d}q_k + i\frac{\partial \mathrm{f}_i}{\partial q_k}\mathrm{d}q_k + j\frac{\partial \mathrm{f}_j}{\partial q_k}\mathrm{d}q_k + k\frac{\partial \mathrm{f}_k}{\partial q_k}\mathrm{d}q_k.
\end{aligned} \tag{2.14}$$

Exploiting the expressions in (2.8) yields

$$\begin{aligned} \mathrm{d}q_r =& \frac{1}{4}\left(\mathrm{d}q + \mathrm{d}q^i + \mathrm{d}q^j + \mathrm{d}q^k\right) & \mathrm{d}q_i =& \frac{1}{4i}\left(\mathrm{d}q + \mathrm{d}q^i - \mathrm{d}q^j - \mathrm{d}q^k\right) \\ \mathrm{d}q_j =& \frac{1}{4j}\left(\mathrm{d}q - \mathrm{d}q^i + \mathrm{d}q^j - \mathrm{d}q^k\right) & \mathrm{d}q_k =& \frac{1}{4k}\left(\mathrm{d}q - \mathrm{d}q^i - \mathrm{d}q^j + \mathrm{d}q^k\right) \end{aligned} \tag{2.15}$$

where substituting (2.15) into (2.14) and after some tedious mathematical manipulations we have

$$\begin{aligned} \mathrm{d}g =& \frac{1}{4}\left(\frac{\partial g}{\partial q_r} - i\frac{\partial g}{\partial q_i} - j\frac{\partial g}{\partial q_j} - k\frac{\partial g}{\partial q_k}\right)\mathrm{d}q + \frac{1}{4}\left(\frac{\partial g}{\partial q_r} - i\frac{\partial g}{\partial q_i} + j\frac{\partial g}{\partial q_j} + k\frac{\partial g}{\partial q_k}\right)\mathrm{d}q^i \\ & \frac{1}{4}\left(\frac{\partial g}{\partial q_r} + i\frac{\partial g}{\partial q_i} - j\frac{\partial g}{\partial q_j} + k\frac{\partial g}{\partial q_k}\right)\mathrm{d}q^j + \frac{1}{4}\left(\frac{\partial g}{\partial q_r} + i\frac{\partial g}{\partial q_i} + j\frac{\partial g}{\partial q_j} - k\frac{\partial g}{\partial q_k}\right)\mathrm{d}q^k. \end{aligned} \tag{2.16}$$

Now, through considering q, q^i, q^j, and q^k as "*algebraically*" independent variables[3], the total differential of $f(q_r, q_i, q_j, q_k) \triangleq f(q, q^i, q^j, q^k)$ is given by

$$\mathrm{d}f = \frac{\partial f}{\partial q}\mathrm{d}q + \frac{\partial f}{\partial q^i}\mathrm{d}q^i + \frac{\partial f}{\partial q^j}\mathrm{d}q^j + \frac{\partial f}{\partial q^k}\mathrm{d}q^k \tag{2.17}$$

where by comparing the expressions in (2.17) and (2.16) it follows that

$$\begin{bmatrix} \frac{\partial f(\mathbf{q},\mathbf{q}^i,\mathbf{q}^j,\mathbf{q}^k)}{\partial \mathbf{q}} \\ \frac{\partial f(\mathbf{q},\mathbf{q}^i,\mathbf{q}^j,\mathbf{q}^k)}{\partial \mathbf{q}^i} \\ \frac{\partial f(\mathbf{q},\mathbf{q}^i,\mathbf{q}^j,\mathbf{q}^k)}{\partial \mathbf{q}^j} \\ \frac{\partial f(\mathbf{q},\mathbf{q}^i,\mathbf{q}^j,\mathbf{q}^k)}{\partial \mathbf{q}^k} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & -i & -j & -k \\ 1 & -i & j & k \\ 1 & i & -j & k \\ 1 & i & j & -k \end{bmatrix}\begin{bmatrix} \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_r} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_i} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_j} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_k} \end{bmatrix} \tag{2.18}$$

which are referred to as the $\mathbb{HR}$-derivatives. In addition, the real-valued components of $q \in \mathbb{H}$ can alternatively be expressed as

$$\begin{aligned} q_r =& \frac{1}{4}\left(q^* + q^{i*} + q^{j*} + q^{k*}\right) & q_i =& \frac{1}{4i}\left(-q^* - q^{i*} + q^j + q^{k*}\right) \\ q_j =& \frac{1}{4j}\left(-q^* + q^{i*} - q^{j*} + q^{k*}\right) & q_k =& \frac{1}{4k}\left(-q^* + q^{i*} + q^{j*} - q^{k*}\right) \end{aligned}$$

resulting in

$$\begin{aligned} \mathrm{d}q_r =& \frac{1}{4}\left(\mathrm{d}q^* + \mathrm{d}q^{i*} + \mathrm{d}q^{j*} + \mathrm{d}q^{k*}\right) & q_i =& \frac{1}{4i}\left(-\mathrm{d}q^* - \mathrm{d}q^{i*} + \mathrm{d}q^j + \mathrm{d}q^{k*}\right) \\ \mathrm{d}q_j =& \frac{1}{4j}\left(-\mathrm{d}q^* + \mathrm{d}q^{i*} - \mathrm{d}q^{j*} + \mathrm{d}q^{k*}\right) & q_k =& \frac{1}{4k}\left(-\mathrm{d}q^* + \mathrm{d}q^{i*} + \mathrm{d}q^{j*} - \mathrm{d}q^{k*}\right) \end{aligned}$$

[3]Note that although q, q^i, q^j, and q^k are essentially the same variable, similar to the format in the $\mathbb{CR}$-calculus, they are considered as "*algebraically*" independent in the context of the $\mathbb{HR}$-calculus as a method for accommodating for the four degrees of freedom inherent in quaternion numbers.

where following the same procedure described for the derivation of the $\mathbb{HR}$-derivatives, leads to the $\mathbb{HR}^*$-derivatives that are given by

$$\begin{bmatrix} \frac{\partial f(\mathbf{q}^*,\mathbf{q}^{i*},\mathbf{q}^{j*},\mathbf{q}^{k*})}{\partial \mathbf{q}^*} \\ \frac{\partial f(\mathbf{q}^*,\mathbf{q}^{i*},\mathbf{q}^{j*},\mathbf{q}^{k*})}{\partial \mathbf{q}^{i*}} \\ \frac{\partial f(\mathbf{q}^*,\mathbf{q}^{i*},\mathbf{q}^{j*},\mathbf{q}^{k*})}{\partial \mathbf{q}^{j*}} \\ \frac{\partial f(\mathbf{q}^*,\mathbf{q}^{i*},\mathbf{q}^{j*},\mathbf{q}^{k*})}{\partial \mathbf{q}^{k*}} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & i & j & k \\ 1 & i & -j & -k \\ 1 & -i & j & -k \\ 1 & -i & -j & k \end{bmatrix} \begin{bmatrix} \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_r} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_i} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_j} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_k} \end{bmatrix}. \tag{2.19}$$

It is important to note that due to the non-commutative nature of quaternions, the placement of the three imaginary unit vectors (i, j, and k) has been consistently kept to the left hand side and hence the expressions in (2.18) and (2.19) are referred to as the left derivatives. Similar results can be obtained by placing the three imaginary unit vectors on the right hand side, known as the right derivatives [85].

Out of all the derivatives of f, of particular interest to signal processing applications is the conjugate derivative given by

$$\frac{\partial f}{\partial q^*} = \frac{1}{4}\left(\frac{\partial f}{\partial q_r} + i\frac{\partial f}{\partial q_i} + j\frac{\partial f}{\partial q_j} + k\frac{\partial f}{\partial q_k}\right)$$

which indicates the direction of the maximum rate of change in f and therefore presenting the $\mathbb{HR}^*$-derivatives as the gradient operator that is [85].

$$\nabla_{\mathbf{q}^{a*}} f = \begin{bmatrix} \frac{\partial f(\mathbf{q},\mathbf{q}^{i},\mathbf{q}^{j},\mathbf{q}^{k})}{\partial \mathbf{q}^*} \\ \frac{\partial f(\mathbf{q},\mathbf{q}^{i},\mathbf{q}^{j},\mathbf{q}^{k})}{\partial \mathbf{q}^{i*}} \\ \frac{\partial f(\mathbf{q},\mathbf{q}^{i},\mathbf{q}^{j},\mathbf{q}^{k})}{\partial \mathbf{q}^{j*}} \\ \frac{\partial f(\mathbf{q},\mathbf{q}^{i},\mathbf{q}^{j},\mathbf{q}^{k})}{\partial \mathbf{q}^{k*}} \end{bmatrix} = \frac{1}{4}\begin{bmatrix} 1 & i & j & k \\ 1 & i & -j & -k \\ 1 & -i & j & -k \\ 1 & -i & -j & k \end{bmatrix} \left(\begin{bmatrix} \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_r} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_i} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_j} \\ \frac{\partial f(\mathbf{q}_r,\mathbf{q}_i,\mathbf{q}_j,\mathbf{q}_k)}{\partial \mathbf{q}_k} \end{bmatrix} = \nabla f \right).$$

The gradient operator can be applied to convex real-valued functions of quaternion-valued variables to find the direction of steepest-descent and to calculate the first-order Taylor series expansion of a quaternion-valued function at $q = q_0 \in \mathbb{H}$ given by [85, 87]

$$f(q) \approx f(q_0) + \Delta\mathbf{q}^{aH}\left(\nabla_{\mathbf{q}^{a*}} f\big|_{\mathbf{q}^a=\mathbf{q}_0^a}\right)$$

where $\Delta\mathbf{q}^{aH} = \mathbf{q}^{aH} - \mathbf{q}_0^{aH}$.

2.7 The augmented quaternion statistics

Similar to the case of complex-valued random variables, which were discussed in Section 2.3, the standard covariance of a quaternion-valued random variable, $\mathbf{q}$, given by $\mathbf{C}_{\mathbf{q}} = E[\mathbf{q}\mathbf{q}^H]$, does not reveal any information on the dependence structure or power differences between the real-valued components of $\mathbf{q}$ and thus only partially describes the second-order statistical information of $\mathbf{q}$. A full description of the second-order statistics of a general quaternion-valued random variable is only possible through the augmented quaternion statistics [71]. In order to account for the entire second-order information of quaternion-valued random variables, a one-to-one relation is established between the quaternion-valued random variable and the vector of its real-valued components, $[\mathbf{q}_r^T, \mathbf{q}_i^T, \mathbf{q}_j^T, \mathbf{q}_k^T]^T$, through exploiting the expressions in (2.8), that is given by [71]

$$\underbrace{\begin{bmatrix} \mathbf{q} \\ \mathbf{q}^i \\ \mathbf{q}^j \\ \mathbf{q}^k \end{bmatrix}}_{\mathbf{q}^a} = \underbrace{\begin{bmatrix} \mathbf{I} & i\mathbf{I} & j\mathbf{I} & k\mathbf{I} \\ \mathbf{I} & i\mathbf{I} & -j\mathbf{I} & -k\mathbf{I} \\ \mathbf{I} & -i\mathbf{I} & j\mathbf{I} & -k\mathbf{I} \\ \mathbf{I} & -i\mathbf{I} & -j\mathbf{I} & k\mathbf{I} \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} \mathbf{q}_r \\ \mathbf{q}_i \\ \mathbf{q}_j \\ \mathbf{q}_k \end{bmatrix} \tag{2.20}$$

where due to the fact that $\mathbf{q}$ has been augmented with its involutions around the i, j, and k axes, $\mathbf{q}^a$ is referred to as the augmented quaternion vector, while $\mathbf{A}$ provides a mapping from $\mathbb{R}^{4N}$ to $\mathbb{H}^{4N}$ the inverse of which is given by $\mathbf{A}^{-1} = \frac{1}{4}\mathbf{A}^H$. The augmented quaternion covariance matrix now becomes

$$\mathbf{C}_{\mathbf{q}^a} = E[\mathbf{q}^a\mathbf{q}^{aH}] = \begin{bmatrix} \mathbf{C}_{\mathbf{q}} & \mathbf{R}_{\mathbf{q}\mathbf{q}^{i*}} & \mathbf{R}_{\mathbf{q}\mathbf{q}^{j*}} & \mathbf{R}_{\mathbf{q}\mathbf{q}^{k*}} \\ \mathbf{R}_{\mathbf{q}^i\mathbf{q}^*} & \mathbf{C}_{\mathbf{q}^i} & \mathbf{R}_{\mathbf{q}^i\mathbf{q}^{j*}} & \mathbf{R}_{\mathbf{q}^i\mathbf{q}^{k*}} \\ \mathbf{R}_{\mathbf{q}^j\mathbf{q}^*} & \mathbf{R}_{\mathbf{q}^j\mathbf{q}^{i*}} & \mathbf{C}_{\mathbf{q}^j} & \mathbf{R}_{\mathbf{q}^j\mathbf{q}^{k*}} \\ \mathbf{R}_{\mathbf{q}^k\mathbf{q}^*} & \mathbf{R}_{\mathbf{q}^k\mathbf{q}^{i*}} & \mathbf{R}_{\mathbf{q}^k\mathbf{q}^{j*}} & \mathbf{C}_{\mathbf{q}^k} \end{bmatrix} \tag{2.21}$$

where $\forall\zeta, \zeta' \in \{1, i, j, k\}, \mathbf{R}_{\mathbf{q}^{\zeta'}\mathbf{q}^{\zeta*}} = E[\mathbf{q}^{\zeta'}\mathbf{q}^{\zeta H}] = \mathbf{R}^H_{\mathbf{q}^{\zeta}\mathbf{q}^{\zeta'*}}$ and $\mathbf{C}_{\mathbf{q}^\zeta} = E[\mathbf{q}^\zeta\mathbf{q}^{\zeta H}]$. It is also straightforward to prove that $\mathbf{C}_{\mathbf{q}^\zeta} = \mathbf{C}^H_{\mathbf{q}^\zeta}$; moreover, it can be shown that $\mathbf{R}_{\mathbf{q}^i\mathbf{q}^{j*}} = \mathbf{R}^i_{\mathbf{q}\mathbf{q}^{k*}}$, $\mathbf{R}_{\mathbf{q}^i\mathbf{q}^{k*}} = \mathbf{R}^i_{\mathbf{q}\mathbf{q}^{j*}}$, and $\mathbf{R}_{\mathbf{q}^j\mathbf{q}^{k*}} = \mathbf{R}^j_{\mathbf{q}\mathbf{q}^{i*}}$. Therefore, the augmented covariance matrix is a Hermitian matrix and can also be presented as

$$\mathbf{C}_{\mathbf{q}^a} = \begin{bmatrix} \mathbf{C}_{\mathbf{q}} & \mathbf{R}_{\mathbf{q}\mathbf{q}^{i*}} & \mathbf{R}_{\mathbf{q}\mathbf{q}^{j*}} & \mathbf{R}_{\mathbf{q}\mathbf{q}^{k*}} \\ \mathbf{R}^H_{\mathbf{q}\mathbf{q}^{i*}} & \mathbf{C}_{\mathbf{q}^i} & \mathbf{R}^i_{\mathbf{q}\mathbf{q}^{k*}} & \mathbf{R}^i_{\mathbf{q}\mathbf{q}^{j*}} \\ \mathbf{R}^H_{\mathbf{q}\mathbf{q}^{j*}} & \mathbf{R}^{iH}_{\mathbf{q}\mathbf{q}^{k*}} & \mathbf{C}_{\mathbf{q}^j} & \mathbf{R}^j_{\mathbf{q}\mathbf{q}^{i*}} \\ \mathbf{R}^H_{\mathbf{q}\mathbf{q}^{k*}} & \mathbf{R}^{iH}_{\mathbf{q}\mathbf{q}^{j*}} & \mathbf{R}^{jH}_{\mathbf{q}\mathbf{q}^{i*}} & \mathbf{C}_{\mathbf{q}^k} \end{bmatrix}. \tag{2.22}$$

Notice that the complete second-order information within the augmented covariance matrix is contained in the standard covariance, $\mathbf{C}_{\mathbf{q}}$, the i-pseudo-covariance, $\mathbf{R}_{\mathbf{q}\mathbf{q}^{i*}}$, the j-pseudo-covariance, $\mathbf{R}_{\mathbf{q}\mathbf{q}^{j*}}$, and the k-pseudo-covariance $\mathbf{R}_{\mathbf{q}\mathbf{q}^{k*}}$. It is now clear that the standard covariance can only fully describe the second-order statistics of a quaternion random variable with vanishing pseudo-covariances referred to as second-order proper. Furthermore, a quaternion random variable is referred to as second-order i-proper if it has vanishing j and k-pseudo-covariances. The second-order j and k-properness are defined in an analogous way.

Now, consider the MMSE estimation of the quaternion-valued variable, y, conditional to the quaternion-valued observation, x. Akin to what was stated for complex-valued random variables in Section 2.3, for quaternion-valued random variables the MMSE estimator has to be expressed according to the real-valued individual components of x and y; therefore, the MMSE estimator is given by

$$\begin{aligned}\hat{y} =& \underbrace{E[y_r|x_r, x_i, x_j, x_k]}_{\hat{y}_r} + i\, \underbrace{E[y_i|x_r, x_i, x_j, x_k]}_{\hat{y}_i} \\ &+ j\, \underbrace{E[y_j|x_r, x_i, x_j, x_k]}_{\hat{y}_j} + k\, \underbrace{E[y_k|x_r, x_i, x_j, x_k]}_{\hat{y}_k}.\end{aligned}$$

The mapping in (2.20) is now exploited to replace the real-valued components of x using the quaternion involutions resulting in

$$\begin{aligned}\hat{y} =& E[y_r|x, x^i, x^j, x^k] + iE[y_i|x, x^i, x^j, x^k] \\ &+ jE[y_j|x, x^i, x^j, x^k] + kE[y_k|x, x^i, x^j, x^k].\end{aligned}$$

Therefore, for quaternion-valued, zero-mean, and jointly Gaussian x and y, the solution is now given in the form of the widely-linear estimator

$$\hat{y} = \mathbf{h}^H\mathbf{x} + \mathbf{g}^H\mathbf{x}^i + \mathbf{u}^H\mathbf{x}^j + \mathbf{v}^H\mathbf{x}^k = \mathbf{w}^{aH}\mathbf{x}^a \tag{2.23}$$

where $\mathbf{h}$, $\mathbf{g}$, $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}^{aH} = [\mathbf{h}^H, \mathbf{g}^H, \mathbf{u}^H, \mathbf{v}^H]$ are vectors of quaternion-valued coefficients, while $\mathbf{x}$ is the regressor vector. The optimal augmented weight vector, $\mathbf{w}^a_{opt}$, that minimizes $E\left[|y-\hat{y}|^2\right]$ is now given by $\mathbf{w}^a_{opt} = \left(E\left[\mathbf{x}^a\mathbf{x}^{aH}\right]\right)^{-1} E\left[\mathbf{x}^a y^*\right]$ [71, 77, 78] and referred to as the Wiener solution.

2.8 The widely-linear quaternion LMS algorithm

A widely-linear quaternion least mean square (WL-QLMS) algorithm for adaptive filtering of quaternion-valued signals has been introduced by Took and Mandic in [70, 88]. This has been achieved by considering the widely-linear model in (2.23) in its adaptive formulation, given by

$$\hat{y}_n = \mathbf{h}_n^H\mathbf{x}_n + \mathbf{g}_n^H\mathbf{x}_n^i + \mathbf{u}_n^H\mathbf{x}_n^j + \mathbf{v}_n^H\mathbf{x}_n^k$$

where the coefficients are updated at each time instant in a steepest-descent fashion according to the gradient of the cost function $J_n = |\varepsilon_n|^2 = \varepsilon_n\varepsilon_n^*$, where $\varepsilon_n = \hat{y}_n - y_n$. The gradient of J_n with respect to $\mathbf{h}$ is calculated through the $\mathbb{HR}$-calculus, as J_n does not admit the Cauchy-Riemann-Fueter conditions, and is given by

$$\nabla_{\mathbf{h}^*} J_n = \left(\nabla_{\mathbf{h}^*}\varepsilon_n\right)\varepsilon_n^* + \varepsilon_n\left(\nabla_{\mathbf{h}_n^*}\varepsilon_n^*\right)$$

with $\nabla_{\mathbf{h}_n^*}\varepsilon_n = \mathbf{x}_n$, whereas using the expression in (2.10) yields

$$\nabla_{\mathbf{h}_n^*}\varepsilon_n^* = \nabla_{\mathbf{h}_n^*}\left(\mathbf{x}_n^H\mathbf{h}_n\right) = \nabla_{\mathbf{h}_n^*}\left(\frac{\mathbf{x}_n^H\mathbf{h}_n^{*i}}{2} + \frac{\mathbf{x}_n^H\mathbf{h}_n^{*j}}{2} + \frac{\mathbf{x}_n^H\mathbf{h}_n^{*k}}{2} - \frac{\mathbf{x}_n^H\mathbf{h}_n^{*}}{2}\right) = \frac{-1}{2}\mathbf{x}_n^*.$$

The gradients $\nabla_{\mathbf{g}_n^*}J_n$, $\nabla_{\mathbf{u}_n^*}J_n$, and $\nabla_{\mathbf{v}_n^*}J_n$ can be calculated in a similar fashion and therefore the updates of the coefficient vectors are given by [70, 88]

$$\begin{aligned}
\mathbf{h}_{n+1} =& \mathbf{h}_n - \mu\left(\mathbf{x}_n\varepsilon_n^* - \frac{1}{2}\varepsilon_n\mathbf{x}_n^*\right) & \mathbf{g}_{n+1} =& \mathbf{g}_n - \mu\left(\mathbf{x}_n^i\varepsilon_n^* - \frac{1}{2}\varepsilon_n\mathbf{x}_n^{i*}\right)\\
\mathbf{u}_{n+1} =& \mathbf{u}_n - \mu\left(\mathbf{x}_n^j\varepsilon_n^* - \frac{1}{2}\varepsilon_n\mathbf{x}_n^{j*}\right) & \mathbf{v}_{n+1} =& \mathbf{v}_n - \mu\left(\mathbf{x}_n^k\varepsilon_n^* - \frac{1}{2}\varepsilon_n\mathbf{x}_n^{k*}\right)
\end{aligned}$$

where $\mu \in \mathbb{R}^+$ denotes the adaptation gain.

Although the WL-QLMS is optimal for both second-order proper and improper quaternion-valued signals, if x and y are jointly proper, then the widely-linear model in (2.23) can be simplified into a strictly linear model given by $\hat{y} = \mathbf{h}^H\mathbf{x}$, mitigating the need for calculating $\mathbf{g}$, $\mathbf{u}$, and $\mathbf{v}$ [77, 78]. In addition, if x and y are jointly i-proper, then the widely-linear model in (2.23) can be simplified into a semi-widely-linear model given by $\hat{y} = \mathbf{h}^H\mathbf{x} + \mathbf{g}^H\mathbf{x}_n^i$, mitigating the need for calculating $\mathbf{u}$ and $\mathbf{v}$ [77, 78], with analogues semi-widely linear models defined for the case where x and y are jointly j and k-proper. Thus, tracking the statistical information in the pseudo-covariance matrices in real-time can lead to significant reduction in computational complexity without loss in performance, not only in the case of WL-QLMS, but also in a variety of quaternion-valued adaptive filtering algorithms [70, 77–79, 85, 87].

2.9 The augmented quaternion Kalman filter

Although a wide range of Kalman filtering algorithms for dealing with quaternion-valued signals have been developed [63, 89–92], the lack of an all inclusive mathematical framework has tied these filtering algorithms to the specific applications that they were designed for. In addition, these Kalman filtering algorithms are not inherently quaternion-valued as in most cases they transform quaternions to their real-valued vector representation and process the signal in $\mathbb{R}^4$. However, the augmented quaternion statistics in conjunction with the $\mathbb{HR}$-calculus have led to the development of a class of quaternion Kalman filters [87] that are suitable for the generality of quaternion signals, are not tied to any specific application, and are directly derived in the quaternion domain.

In a similar fashion to what was described for the AECKF in Section 2.3, consider this time the quaternion-valued augmented state vector evolution sequence $\{\mathbf{x}_n^a, n = 0, 1, 2, \ldots\}$ given by

$$\mathbf{x}_n^a = f_n(\mathbf{x}_{n-1}^a) + \boldsymbol{\nu}_n^a$$

where $f_n(\cdot)$ and $\boldsymbol{\nu}_n^a$ are the state evolution function and augmented state evolution noise at time instant n. Recall that the objective is to track $\mathbf{x}_n^a$ in real-time; however, only noisy observations of the augmented state vector are at hand that can be modeled as

$$\mathbf{y}_n^a = h_n(\mathbf{x}_n^a) + \boldsymbol{\omega}_n^a$$

where $\mathbf{y}_n^a$, $h_n(\cdot)$, and $\boldsymbol{\omega}_n^a$ are the augmented observation vector, the observation function, and the augmented observational noise vector at time instant n.

For the sake of simplicity, hereafter we shall approximate the state evolution and observation functions in a widely-linear fashion as $f_n(\mathbf{x}_n^a) \approx \mathbf{A}_n^a \mathbf{x}_n^a$ and $h_n(\mathbf{x}_n^a) \approx \mathbf{H}_n^a \mathbf{x}_n^a$ through applying the $\mathbb{HR}$-calculus, with $\mathbf{A}_n^a$ and $\mathbf{H}_n^a$ denoting the Jacobian matrices of $f_n(\cdot)$ and $h_n(\cdot)$. Then, the trace of the augmented covariance matrix of the *a posteriori* estimation error vector at time instant n is minimized through implementing the augmented quaternion Kalman filter (AQKF) that is given in its information formulation in Algorithm 2, where $\mathbf{C}_{\boldsymbol{\nu}_n^a}$ and $\mathbf{C}_{\boldsymbol{\omega}_n^a}$ denote the augmented covariance matrices of $\boldsymbol{\nu}_n^a$ and $\boldsymbol{\omega}_n^a$, while $\hat{\mathbf{x}}_{n|n-1}^a$ and $\hat{\mathbf{x}}_{n|n}^a$ represent the *a priori* and *a posteriori* estimates of $\mathbf{x}_n^a$, whereas $\hat{\mathbf{M}}_{n|n-1}^a$ and $\hat{\mathbf{M}}_{n|n}^a$ denote the augmented covariance matrices of the *a priori* and *a posteriori* state estimation error vectors. [85, 87].

Algorithm 2. AQKF [87]

Initialize with:

$$\begin{aligned}\hat{\mathbf{x}}_{0|0}^a &= E[\mathbf{x}_0^a] \\ \hat{\mathbf{M}}_{0|0}^a &= E\big[(\mathbf{x}_0^a - E[\mathbf{x}_0^a])(\mathbf{x}_0^a - E[\mathbf{x}_0^a])^H\big]\end{aligned}$$

Model update:

$$\begin{aligned}\hat{\mathbf{x}}_{n|n-1}^a &= \mathbf{A}_n^a \hat{\mathbf{x}}_{n-1|n-1}^a \\ \hat{\mathbf{M}}_{n|n-1}^a &= \mathbf{A}_n^a \hat{\mathbf{M}}_{n-1|n-1}^a \mathbf{A}_n^{aH} + \mathbf{C}_{\boldsymbol{\nu}_n^a}\end{aligned}$$

Measurement update:

$$\begin{aligned}\hat{\mathbf{M}}_{n|n}^{a^{-1}} &= \hat{\mathbf{M}}_{n|n-1}^{a^{-1}} + \mathbf{H}_n^{aH} \mathbf{C}_{\boldsymbol{\omega}_n^a}^{-1} \mathbf{H}_n^a \\ \mathbf{G}_n^a &= \hat{\mathbf{M}}_{n|n}^a \mathbf{H}_n^{aH} \mathbf{C}_{\boldsymbol{\omega}_n^a}^{-1} \\ \hat{\mathbf{x}}_{n|n}^a &= \hat{\mathbf{x}}_{n|n-1}^a + \mathbf{G}_n^a \big(\mathbf{y}_n^a - \mathbf{H}_n^a \hat{\mathbf{x}}_{n|n-1}^a\big)\end{aligned}$$

From Algorithm 2, notice that the AQKF uses the state and observation vectors in their augmented formulation. In addition, in order to incorporate their full second-order statistical information, the augmented covariance matrix of the state evolution and observation noise vectors are required to implement the AQKF. Although this approach is optimal for the generality of quaternion-valued signals, it might impose an excessive amount of computational burden on the processing unit when handling large state and observation vectors. In order to reduce the computational complexity of the AQKF, the efficient implementation of the AQKF based on the structure of augmented covariance matrices are proposed in [85, 87], where it is also shown that the widely-linear model can be simplified when it comes to processing three-dimensional data modeled as pure quaternions, which can further reduce the computational requirements of the AQKF.

Chapter 3

Real-Time Tracking of Quaternion Impropriety

3.1 Overview

An algorithm for tracking the degree of quaternion impropriety in real-time is developed. This is achieved through exploiting the i, j, and k-pseudo-covariances that make possible the introduction of an impropriety measure as the MMSE solution for estimating the quaternion involutions along the i, j, and k axes from the quaternion random variable itself. The performance of the developed real-time quaternion impropriety tracker both in the mean and mean square error (MSE) sense are analyzed allowing to establish convergence bounds and quantify the effect of the degree of impropriety on the steady-state performance of the developed real-time quaternion impropriety tracker. The concept is verified through simulations on both synthetic and real-world data.

3.2 Introduction

The quaternion widely-linear model is based on augmenting the quaternion random variable with its involutions along the i, j, and k axes [69–71, 77–79]. Therefore, quaternion-valued signal processing algorithms established on the quaternion widely-linear model have four times as many parameters as their strictly linear counterparts. However, as was discussed in Section 2.8, in instances when the signal is proper, widely-linear quaternion-valued signal processing algorithms can be simplified in order to reduce their computational complexity without negatively effecting their performance [77, 78]. In addition to reducing computational complexity, the higher number of updates that have to be calculated in widely-linear algorithms result in a higher gradient noise in gradient-based learning methods and slower convergence rates. Thus, it becomes essential to identify the degree of impropriety of a signal in real-time, in both detection and estimation applications, so that the instants when a non-stationary signal changes its statistic can be identified and an estimator that best suits the signal can be selected.

The properness of complex-valued random variables has been extensively studied [33, 43, 83, 93–95] with an impropriety measure introduced and its effect from a geometric stand point investigated in [93]. However, In contrast to complex-valued random variables, properness of quaternion-valued random variables has not yet been thoroughly addressed. The approach taken in [96] is based on the probability distribution function (pdf) and considers quaternion properness as the invariance of the pdf under specific rotations. The concept was taken further in [97], where the condition for quaternion properness was considered as invariance of the pdf under rotations around any axis and for any angle that is

$$\forall\theta \in [0, 2\pi) \text{ and } \xi \in \mathbb{H} \text{ such that } \xi^2 = -1 \text{ then } P_Q(q) = P_Q\left(e^{\xi\theta} q\right).$$

In [71], a quaternion-valued Gaussian random variable is defined as proper if it has vanishing pseudo-covariances. Three different types of quaternion properness based on vanishing of three different pseudo-covariances were defined and their impact on the quaternion widely-linear model and quaternion-valued signal processing techniques were analyzed in [77, 78]. In addition, an algorithm for measuring each type of impropriety based on the Kullback-Leibler divergence between multivariate quaternion-valued Gaussian distributions has been proposed in [98]. However, an algorithm for tracking quaternion impropriety of non-stationary signals in real-time is still lacking.

In this chapter, we introduce a novel algorithm for real-time tracking of quaternion impropriety based on quaternion-valued adaptive filtering. This is achieved through introducing three impropriety measures established on the i, j, and k-pseudo-covariances and illustrating that each impropriety measure is the MMSE solution for estimating the involutions of a quaternion random variable along the i, j, and k axes from the quaternion random variable itself. For rigor, the performance of the developed impropriety tracker is analyzed in order to establish convergence conditions and quantify the effect of the degree of impropriety on the steady-state performance of the developed impropriety tracker. Finally, the performance of the developed real-time quaternion impropriety tracker is verified through simulations on both synthetically generated and real-world data recordings.

3.3 Quaternion impropriety measures

The properness of a quaternion-valued random variable reflects the dependence structure of its real-valued components, that is the ratio of signal powers and/or correlation between the real-valued components of the quaternion-valued random variable. Therefore, quaternion properness can be related to the properness of the projection of a quaternion-valued random variable on the six complex planes denoted by 1-i, 1-j, 1-k, i-j, i-k, and j-k, where "1" represents the real axis [71, 77, 78]. Thus, measuring the complex impropriety, defined as the ratio between the pseudo-covariance and the covariance, in these six planes reveals the complete dependence structure of the real-valued components of a quaternion-valued random variable and hence provides a measure of its impropriety.

The structure of the quaternion covariance and pseudo-covariances are given in Table 3.1, from which notice that the six complex impropriety measures, corresponding to the six mentioned complex planes, can be extracted from the i, j, and k-pseudo-covariances. Thus, similar to the approaches in [71, 77, 78, 99], we can now define the following three impropriety measures, $\rho_\zeta = \{\rho_i, \rho_j, \rho_k\}$, for quaternion-valued random variables

$$\begin{aligned} \rho_i &= C_q^{-1} R_{qq^{i*}} = \left(E[qq^*]\right)^{-1} E[qq^{i*}] \\ \rho_j &= C_q^{-1} R_{qq^{j*}} = \left(E[qq^*]\right)^{-1} E[qq^{j*}] \\ \rho_k &= C_q^{-1} R_{qq^{k*}} = \left(E[qq^*]\right)^{-1} E[qq^{k*}] \end{aligned} \tag{3.1}$$

which in essence represent the correlation between the quaternion valued random variable q and its involution around $\zeta \in \{i, j, k\}$, normalized by the signal power, $C_q = E[qq^*]$.

Table 3.1: Structure of the quaternion covariance and pseudo-covariances.

	$\Re\{\cdot\}$	$\Im_i\{\cdot\})$	$\Im_j\{\cdot\}$	$\Im_k\{\cdot\}$
C_q	$C_{q_r} + C_{q_i} + C_{q_j} + C_{q_k}$	0	0	0
$R_{qq^{i*}}$	$C_{q_r} + C_{q_i} - C_{q_j} - C_{q_k}$	0	$2(R_{q_r q_j} - R_{q_i q_k})$	$2(R_{q_r q_k} + R_{q_i q_j})$
$R_{qq^{j*}}$	$C_{q_r} - C_{q_i} + C_{q_j} - C_{q_k}$	$2(R_{q_r q_i} + R_{q_j q_k})$	0	$2(R_{q_r q_k} - R_{q_i q_j})$
$R_{qq^{k*}}$	$C_{q_r} - C_{q_i} - C_{q_j} + C_{q_k}$	$2(R_{q_r q_i} - R_{q_j q_k})$	$2(R_{q_r q_j} + R_{q_i q_k})$	0

3.4 Tracking quaternion impropriety in real-time

Consider the problem of finding the optimal linear mapping that relates the quaternion-valued random variable q to its involution q^ζ with $\zeta \in \{i, j, k\}$. This mapping can be formulated as

$$q^\zeta = h_{opt}^* q$$

where applying the conjugate operator and multiplying both sides by q gives

$$qq^{\zeta *} = qq^* h_{opt}. \tag{3.2}$$

The closed form solution for h_{opt} can now be found through taking the statistical expectation of the expression in (3.2) which yields

$$h_{opt} = \left(E[qq^*]\right)^{-1} E[qq^{\zeta *}]. \tag{3.3}$$

Note that the expression in (3.3) is not only the ρ_ζ impropriety measure, but also the MMSE or Wiener solution for estimating the ζ-involution of q form itself, that can be formulated as

$$\hat{q}^\zeta = h^* q$$

which minimizes $E\left[|\varepsilon|^2\right] = E\left[|\hat{q}^\zeta - q|^2\right]$, where $\hat{q}^\zeta$ denotes the estimate of q^ζ. However, finding the Wiener solution requires knowledge of the true statistics of the signal that in general are not available. Moreover, applying block based estimators for finding the Wiener solution is rather inadequate for on-line applications, specially when dealing with non-stationary signals with fast changing statistics, or in applications where it becomes important to capture incidences that the signal changes its statistics. Thus, an adaptive impropriety estimator is required.

The definition of the impropriety measure as the optimal Wiener solution for estimating q^ζ from q permits the application of the strictly linear quaternion least mean square (QLMS) adaptive filter, used here in its iQLMS [100] formulation, to track the impropriety measure in real-time. The operations of such a real-time quaternion impropriety tracker are summarized in Algorithm 3, where $\zeta \in \{i, j, k\}$ and the adaptation gain is denoted by $\mu \in \mathbb{R}^+$. As the QLMS algorithm uses instantaneous estimates of the signal statistics, the filter coefficients never reach their optimal values in the absolute sense and therefore it becomes important to analyze the contribution of the bias and variance of the parameter estimates to the total MSE.

Algorithm 3. Real-Time Quaternion Impropriety Tracker

For each time instant $n = 1, 2, ...$

Estimate the ζ*-involution of* q:

$$\hat{q}_n^\zeta = h_n^* q_n$$

Calculate the estimation error:

$$\varepsilon_n = q_n^\zeta - \hat{q}_n^\zeta$$

Update the impropriety estimate:

$$h_{n+1} = h_n + \frac{\mu}{2} q_n \varepsilon_n^*$$

3.5 Performance analysis

In order to understand the behavior of the proposed quaternion impropriety tracker, in the sequel, we shall analyze the behavior of the weight error given by

$$\epsilon_n = h_n - h_{opt}. \tag{3.4}$$

3.5.1 Mean error behavior

In order to find a bound on the adaptation gain which insures unbiased operation of the developed quaternion impropriety tracker, the statistical expectation of the weigh error at each time instant must be expressed in a recursive fashion. To this end, notice that from Algorithm 3 we have

$$h_{n+1} - h_n = \frac{\mu}{2} q_n \varepsilon_n^*. \tag{3.5}$$

In addition, replacing (3.4) into the error $\varepsilon_n^* = q_n^{\zeta*} - q_n^* h_n$ gives

$$\varepsilon_n^* = q_n^{\zeta*} - q_n^* \epsilon_n + q_n^* h_{opt}. \tag{3.6}$$

Now, substituting (3.6) into (3.5) yields

$$\epsilon_{n+1} - \epsilon_n = \frac{\mu}{2} q_n (q_n^{\zeta*} - q_n^* h_n)$$

that can be rearranged to give

$$\epsilon_{n+1} = \epsilon_n + \frac{\mu}{2} q_n q_n^{\zeta*} - \frac{\mu}{2} q_n q_n^* \epsilon_n - \frac{\mu}{2} q_n q_n^* h_{opt}. \tag{3.7}$$

Taking the statistical expectation of (3.7) and replacing h_{opt} with the expression in (3.3), we arrive at the recursive expression for the statistical expectation of the weight error that is given by

$$E[\epsilon_{n+1}] = E[\epsilon_n]\Big(1 - \frac{\mu}{2} E[q_n q_n^*]\Big).$$

It now becomes apparent that the statistical expectation of the weight error converges to zero for

$$\Big|1 - \frac{\mu}{2} E[q_n q_n^*]\Big| < 1$$

so that the allowable range for the adaptation gain becomes

$$0 < \mu < \frac{4}{E[q_n q_n^*]} = \frac{4}{C_{q_n}}. \tag{3.8}$$

Note that the algorithm will operate in an asymptotically unbiased fashion, if the selected adaption gain satisfies the condition set in (3.8). Furthermore, it is important to note that the convergence condition in (3.8) indicates that convergence in the mean is not influenced by the degree of impropriety of the input signal.

3.5.2 Mean square error behavior

In steady-state operating conditions the variance of the weight error is given by

$$E\big[\epsilon_{n+1}\epsilon_{n+1}^*\big] = E\big[|\epsilon_{n+1}|^2\big] = E\big[|\epsilon_n + \frac{\mu}{2}q_n\varepsilon_n^*|^2\big]$$

where upon substituting $\varepsilon_n = q_n^{\zeta*} - q_n^* h_n = q^{\zeta*} - q^*(\epsilon_n + h_{opt})$ we have

$$E[|\epsilon_{n+1}|^2] = E\Big[\big|\epsilon_n + \frac{\mu}{2}q_n\big(q_n^{\zeta*} - q_n^*(\epsilon_n + h_{opt})\big)\big|^2\Big].$$

Assuming that the adaptation gain meets the condition set in (3.8) and the algorithm converges in the mean, it is reasonable to consider that in the steady-state $h_{n+1} \approx h_n$, so that $E[\epsilon_n] \approx 0$, which yields

$$\begin{aligned} E\big[|\epsilon_{n+1}|^2\big] =& E\big[|\epsilon_n|^2\big]\big(1 + \frac{\mu^2}{4}E\big[|q_nq_n^*|^2\big] - \mu E[q_nq_n^*]\big) \\ &+ \frac{\mu^2}{4}\Big(E\big[|q_nq_n^{\zeta*}|^2\big] + E\big[|q_nq_n^*|^2\big]E\big[|h_{opt}|^2\big]\Big) \\ &- \frac{\mu^2}{2}E\Big[\Re(q_nq_n^{\zeta*}h_{opt}^*q_n^*q_n)\Big]. \end{aligned} \tag{3.9}$$

Finally, in steady-state operating conditions it is also reasonable to assume that $E\big[|\epsilon_{n+1}|^2\big] \approx E\big[|\epsilon_n|^2\big]$, which allows the expression in (3.9) to be simplified into

$$E\big[|\epsilon_n|^2\big] = \frac{\frac{\mu}{4}\chi}{E[q_nq_n^*] - \frac{\mu}{4}E\big[|q_nq_n^*|^2\big]}. \tag{3.10}$$

where

$$\begin{aligned} \chi =& E\big[|q_nq_n^{\zeta*}|^2\big] + E\big[|q_nq_n^*|^2\big]E\big[|h_{opt}|^2\big] - 2E\Big[\Re(q_nq_n^{\zeta*}h_{opt}^*q_n^*q_n)\Big] \\ =& E\Big[\Big|q_nq_n^{\zeta*} - q_nq_n^*h_{opt}\Big|^2\Big] \end{aligned} \tag{3.11}$$

For the algorithm to converge in the mean square sense the steady-state weight error variance, as expressed in (3.10), needs to remain positive and bounded. From (3.11) notice that $\chi \in \mathbb{R}^+$; therefore, $E\big[\,|v_n|^2\,\big]$ is positive and bonded if and only if

$$0 < \mu < 4\frac{E[q_nq_n^*]}{E\big[|q_nq_n^*|^2\big]}. \tag{3.12}$$

Furthermore, in Appendix A it is shown that

$$E\big[|q_nq_n^*|^2\big] = \frac{C_{q_n}^2}{2}\left(3 + |\rho_i|^2 + |\rho_j|^2 + |\rho_k|^2\right). \tag{3.13}$$

Therefore, replacing (3.13) into (3.12) allows the mean square convergence condition to be expressed as

$$0 < \mu < \frac{8}{C_{q_n}\left(3 + |\rho_i|^2 + |\rho_i|^2 + |\rho_k|^2\right)}. \tag{3.14}$$

From the expression obtained for the mean square error in (3.10) and the mean square convergence condition in (3.14), observe that the mean square behavior of the developed quaternion impropriety tracker is dependent on the degree of impropriety, as the terms χ and $E\big[|q_n q_n^*|^2\big]$ contain impropriety information.

3.6 Simulations

In this section the performance of the developed quaternion impropriety tracker is validated through simulation on synthetically generated signals and real-world wind data recordings. In addition, the developed quaternion impropriety tracker is applied to track the degree of channel diversity in communications systems based on Alamouti coding.

3.6.1 Synthetically generated data

First, the impropriety tracking ability of the developed algorithm is demonstrated on a synthetically generated signal constructed from three segments of zero-mean unit power white quaternion-valued Gaussian noises with changing pseudo-covariances, and hence impropriety measures. The developed quaternion impropriety tracker was applied to track the degree of impropriety of the signal with $\mu = 0.1$. The absolute values of the quaternion impropriety measures for the data segments with different improprieties together with their estimates is shown in Figure 3.1. Observe that the developed quaternion impropriety tracker produced accurate impropriety estimations for various types of impropriety. In addition, the degree of impropriety arising from one of the pseudo-covariances did not affect the steady-state performance or convergence of other impropriety measures.

To further illustrate the ability of the proposed impropriety tracker and analyze its mean square error performance, we considered zero-mean unit power white quaternion-valued Gaussian noise with changing ρ_j only. The j-impropriety measure, ρ_j, was set to 0.65 for the first segment, 1 for the second segment, and 0.3 for the third segment. In Figure 3.2, 100 realizations of the estimate of ρ_j and their average are shown, demonstrating that the proposed algorithm produces unbiased estimates and that convergence is not affected by the degree of impropriety, which verifies the analysis in Section 3.5.1 and Section 3.5.2. Moreover, observe that the steady-state variance of the impropriety tracker depends on the degree of impropriety.

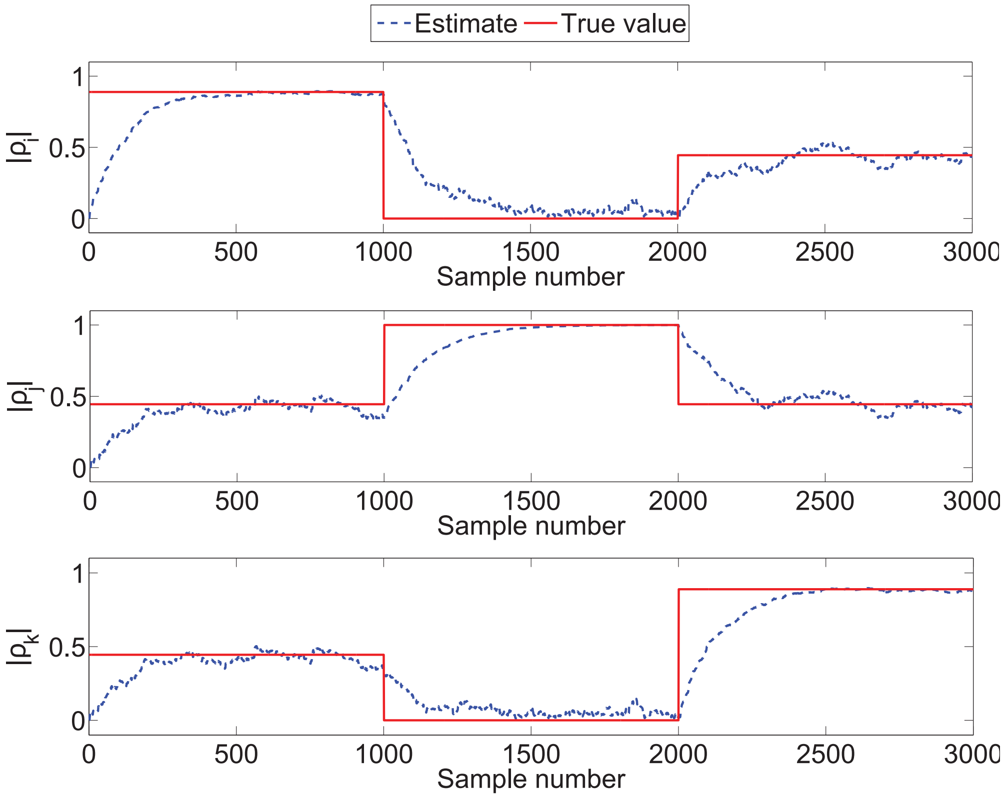

Figure 3.1: Absolute quaternion impropriety measures for synthetically generated Gaussian data (in red) and its quaternion impropriety estimates (in blue).

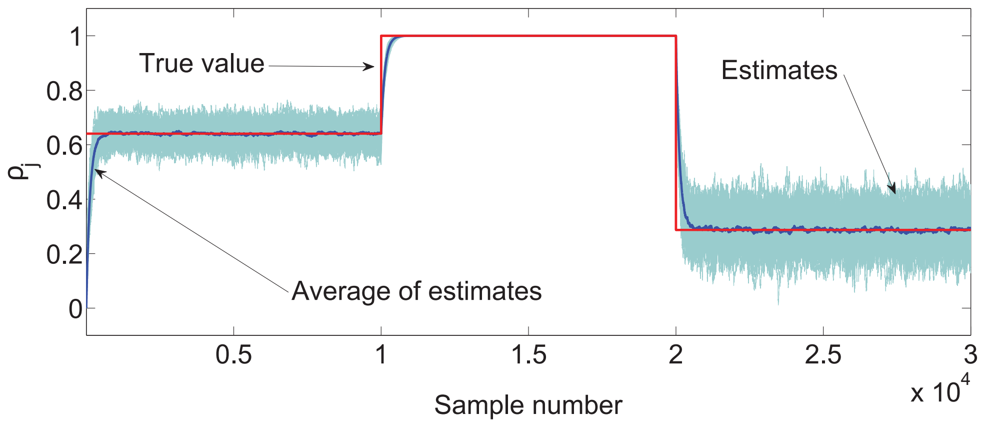

Figure 3.2: True value of ρ_j (in red) plotted alongside 100 realizations of its estimate (in light green) and the average of the estimates (in blue).

3.6.2 Real-world wind data

The quaternion-valued wind data comprised of the wind speed measured in the north, east, and vertical directions as the pure quaternion part and the ambient temperature as the real part[1]. The recorded wind signal exhibits a high degree of impropriety, as

[1]Note that the same representation for wind data has been successfully used in [75, 79] for wind speed and atmospheric temperature prediction.

seen in the scatter diagram in Figure 3.3. Figure 3.4 shows the impropriety measures of the recorded wind signal with the adaptation gain set to $\mu = 0.1$.

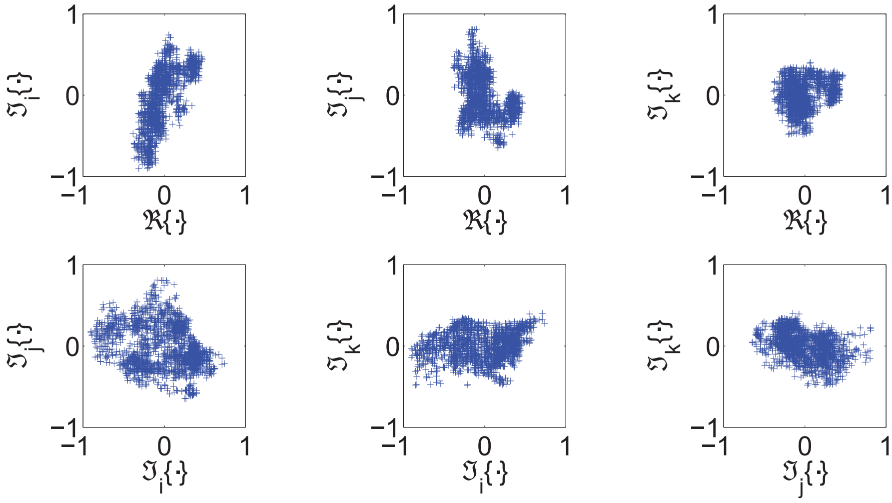

Figure 3.3: Scatter diagram of the improper distribution of wind data.

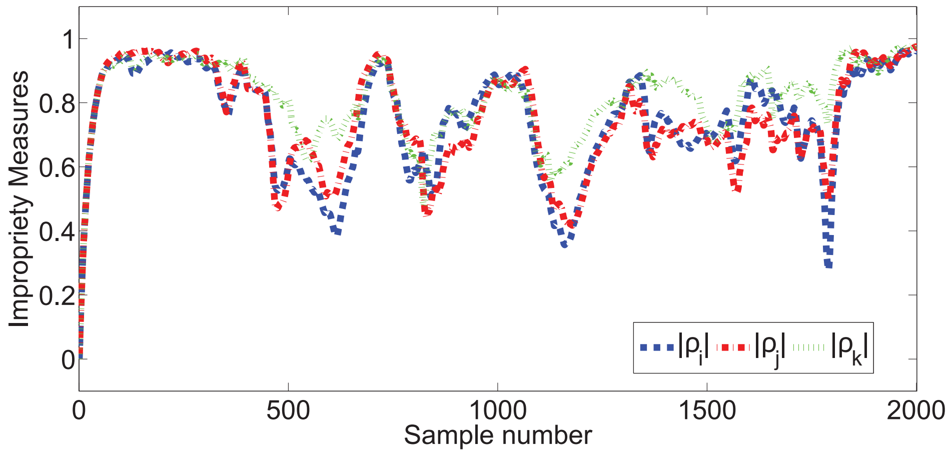

Figure 3.4: Absolute value of the estimated impropriety measures of quaternion-valued wind data.

3.6.3 Communication channel estimation

A multiple-input-multiple-output wireless communication system based on Alamouti coding [101] was considered, where the coding scheme is given by

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_1 & -h_2^* \\ h_2 & h_1^* \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} \tag{3.15}$$

while y_1 and y_2 are two consecutive complex-valued received signals, s_1 and s_2 are two consecutive complex-valued transmitted signals, h_1 and h_2 are complex-valued channel gains between each transmit antenna and the receiver, whereas ω_1 and ω_2 represent complex-valued noise terms.

Using the Cayley-Dickson representation, two complex numbers can be combined into a quaternion, giving the quaternion form of the Alamouti code as [99, 102]

$$Y = HS + W$$

where $Y = y_1 + y_2 j$, $H = h_1 + h_2 j$, $S = s_1 + s_2 j$, and $W = \omega_1 + \omega_2 j$. The impropriety of the channel can be used to analyze its diversity and to establish whether any phase information can be extracted from the received signal [103]. To illustrate this point, the structure of the i, j, and k-pseudo-covariances of H are given in Table 3.2. Note that the entire joint second-order statistical information of h_1 and h_2 can be extracted from pseudo-covariances of H.

Table 3.2: Structure of the pseudo-covariances of the quaternion-valued Alamouti communication channel.

	1-i component	j-k component
$R_{HH^{i*}}$	$C_{h_1} - C_{h_2}$	$2R_{h_2h_1}j$
$R_{HH^{j*}}$	$R_{h_1h_1} + R_{h_2h_2}$	$\left(R_{h_2h_1^*} - R_{h_1h_2^*}\right)j$
$R_{HH^{k*}}$	$R_{h_1h_1} - R_{h_2h_2}$	$\left(R_{h_1h_2^*} + R_{h_2h_1^*}\right)j$

For the first segment (0 to 2.5 seconds) the two complex valued channels in (3.15) were independent, one was circular complex and the other was improper complex with a complex impropriety measure of 0.8. For the second segment (2.5 to 5 seconds) both channels were complex circular and had a cross-correlation of 0.4. Figure 3.5 illustrates the ability of the proposed impropriety tracker to successfully track the changes in channel statistics, where the adaptation gain was $\mu = 0.01$ and the channel was measured every 5 milliseconds.

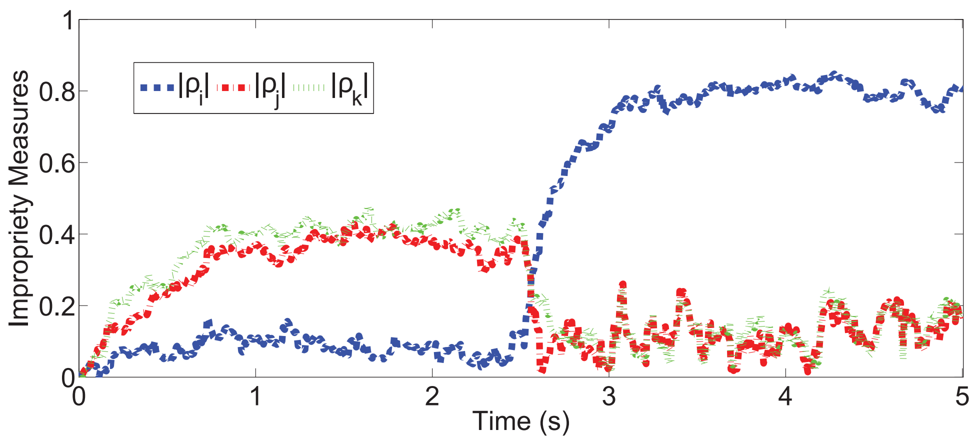

Figure 3.5: Channel impropriety measure estimation of an Alamouti communication system.

3.7 Conclusion

A real-time tracker of quaternion impropriety has been introduced. This has been achieved based on the MMSE linear estimation of the involutions of a quaternion random variable along the i, j, and k-axis from the quaternion random variable itself. Convergence conditions in the mean and mean square sense have also been obtained. The analysis has shown that the proposed algorithm produces unbiased estimates and that the mean behavior of the algorithm is not affected by the degree of impropriety. However, the steady-state variance of the proposed algorithm does exhibit strong dependence on the degree of impropriety. The analysis has been verified using simulations on both synthetic and real-world data.

Chapter 4

A Quaternion Adaptive Phase-Only Estimator

4.1 Overview

Quaternions have proven to be advantageous for modeling three-dimensional rotations in a number of applications and are now considered a favored alternative to rotation matrices. However, despite their natural ability to model phase, when it comes to adaptive phase-only estimation tasks, quaternions remain underutilized. This issue is addressed in this chapter through the introduction of the widely-linear quaternion least mean phase (WL-QLMP) algorithm and its strictly linear counterpart, the quaternion least mean phase (QLMP), for adaptive phase-only estimation of quaternion-valued signals. This is achieved through the derivation of an adaptive phase-only estimator that updates the weights of the adaptive filter at each time instant according to a cost function of the phase error in a steepest-descent fashion based on the $\mathbb{HR}$-calculus. A quantitative assessment of the performance of the developed algorithm is conducted, the physical interpretation of the operations of developed algorithm is provided, and the concept is validated in a number of practical applications including body-motion tracking and the estimation of the fundamental frequency of a three-phase power system under different operating conditions.

4.2 Introduction

The WL-QLMS and its strictly linear dual are the two initial algorithms in quaternion-valued signal processing, which are based on the minimization of the MSE, a real-valued function of quaternion variables [69, 70, 75, 79, 85]. However, in many applications involving three and four-dimensional signals, phase is particularly important as the amplitude information could be corrupted, such as in cases when the signal experiences real-valued multiplicative noise. Moreover, the amplitude information may not be essential, such as when modeling three-dimensional rotations. To illustrate this point, recall from Section 2.5, that if the pre-rotation Cartesian coordinates, $\mathbf{u} = [u_x, u_y, u_z]^T$, is modeled as a pure quaternion given by $q_{\mathbf{u}} = iu_x + ju_y + ku_z$, then the post-rotation coordinates of $q_{\mathbf{u}}$ by an angle of θ around the unit vector $\boldsymbol{\eta}$ can be calculated through the involution

$$q_{\mathbf{u}'} = \underbrace{\left(\cos\left(\frac{\theta}{2}\right) + q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\right)}_{\varsigma} q_{\mathbf{u}} \underbrace{\left(\cos\left(\frac{\theta}{2}\right) - q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\right)}_{\varsigma^{-1}} \tag{4.1}$$

where $q_{\mathbf{u}'} = iu'_x + ju'_y + ku'_z$ represents the post-rotation coordinates, while $q_{\boldsymbol{\eta}} = i\eta_x + j\eta_y + k\eta_z$. Note that the amplitude of ς has no bearing on the outcome of the rotation expressed in (4.1) as the only parameters defining the rotation are θ and $\boldsymbol{\eta}$. In such scenarios, the optimization task is to minimize a measure of the phase error, indeed amplitude variations may even have a detrimental effect on the performance. However, such phase-only estimation algorithms for quaternion-valued signals are still lacking.

The direct application of the model in (4.1) results in a non-linear estimator; however, replacing ς^{-1} with its real-valued components gives

$$q_{\mathbf{u}'} = \varsigma q_{\mathbf{u}} \Re(\varsigma^{-1}) + \varsigma q_{\mathbf{u}} \Im_i(\varsigma^{-1}) + \varsigma q_{\mathbf{u}} \Im_j(\varsigma^{-1}) + \varsigma q_{\mathbf{u}} \Im_k(\varsigma^{-1})$$

which can be rearranged into a more elegant representation given by

$$q_{\mathbf{u}'} = \varsigma\left[\Re(\varsigma^{-1}), \Im_i(\varsigma^{-1}), \Im_j(\varsigma^{-1}), \Im_k(\varsigma^{-1})\right]^T \mathbf{q}^a_{\mathbf{u}} = \boldsymbol{\varsigma}^{aT} \mathbf{q}^a_{\mathbf{u}} \tag{4.2}$$

where $\boldsymbol{\varsigma}^a = \varsigma\left[\Re(\varsigma^{-1}), \Im_i(\varsigma^{-1}), \Im_j(\varsigma^{-1}), \Im_k(\varsigma^{-1})\right]^T$, a form consistent with the widely-linear model (see Section 2.7).

In the complex domain, phase estimation is preformed by the least mean phase (LMP) [104] and the least mean magnitude phase (LMMP) [105] adaptive filtering algorithms. The LMP algorithm employs the phase error cost function, and has shown superior performance compared to the CLMS algorithm when used for channel equalization in communications applications [104]. The LMMP algorithm decomposes the MSE into the amplitude and phase errors, and was implemented for channel equalization in the presence of Doppler shift induced by physical motion and in array processing, outperforming the CLMS algorithm in both applications [105]. Based on complex-valued widely-linear modeling, the widely-linear least mean phase (WL-LMP) algorithm was proposed in [56] for real-time estimation of the fundamental frequency in a three-phase power system, where it outperformed conventional frequency estimation methods for unbalanced power systems. The WL-LMP algorithm was also shown to be less sensitive to amplitude variations, a desirable property of phase-only estimators.

In this chapter, the complex-valued LMP algorithm is generalized to the quaternion domain and the widely-linear setting, in order to provide a rigorous quaternion adaptive phase-only estimator. To this end, the WL-QLMP adaptive filtering algorithm for both second-order proper and improper quaternion-valued processes is introduced. In addition, a convenient geometric interpretation and stability analysis is also provided. The performance of the proposed WL-QLMP and QLMP are validated over two practical case studies, tracking the rotation of the limbs of an athlete while performing Tai-Chi movements, and in estimating the fundamental frequency of a three-phase power system under different operating conditions.

4.3 The quaternion least mean phase estimator

The polar representation of a quaternion number is given by $q = |q|e^{\xi\theta}$, where ξ is the normalized projection of q onto the imaginary subspace of $\mathbb{H}$. Given that ξ and the real axis are orthogonal and $\xi^2 = -1$, the real axis along with ξ define a two-dimensional plane in $\mathbb{H}$, which is isomorphic to the complex domain and includes q. In this subspace, θ is the angle between the real axis and q. Consequently, $\xi\theta$ uniquely describes the orientation of q; thus, we consider $\xi\theta$ as the phase of the quaternion variable q.

Recall from Section 2.7, that the quaternion widely-linear estimator allows us to estimate the process y through a widely-linear estimator $\hat{y} = \mathbf{w}^{aH}\mathbf{q}^a$, where $\mathbf{w}^a$ and $\mathbf{q}^a$ are respectively the weight vector and augmented observation or regressor vector. The phase error of the estimation at a time instant n is then given by

$$\varepsilon_n = \xi_{\hat{y}_n}\theta_{\hat{y}_n} - \xi_{y_n}\theta_{y_n}. \tag{4.3}$$

The phase-only cost function now becomes

$$J_n = \varepsilon_n\varepsilon_n^* = |\xi_{\hat{y}_n}\theta_{\hat{y}_n} - \xi_{y_n}\theta_{y_n}|^2 \tag{4.4}$$

while the steepest-descent weight vector update is given by

$$\mathbf{w}_{n+1}^a = \mathbf{w}_n^a - \mu\nabla_{\mathbf{w}_n^{a*}} J_n \tag{4.5}$$

where $\mu \in \mathbb{R}^+$ denotes the adaptation gain.

In order to be able to implement the proposed quaternion phase estimator the gradient of the cost function, $\nabla_{\mathbf{w}_n^{a*}} J_n$, is calculated next, where for the sake of simplicity in presentation, time indices are dropped. From (4.3), observe that the phase error is a pure quaternion; therefore, $\varepsilon = -\varepsilon^*$ and $J = -\varepsilon^2$. As the cost function is a real-valued function of quaternion variables, by employing the $\mathbb{HR}$-calculus the gradient of the cost function becomes

$$\nabla_{\mathbf{w}^{a*}} J = -(\nabla_{\mathbf{w}^{a*}} \varepsilon)\varepsilon - \varepsilon(\nabla_{\mathbf{w}^{a*}} \varepsilon) \tag{4.6}$$

where

$$\nabla_{\mathbf{w}^{a*}} \varepsilon = (\nabla_{\mathbf{w}^{a*}} \xi_{\hat{y}})\theta_{\hat{y}} + \xi_{\hat{y}}(\nabla_{\mathbf{w}^{a*}} \theta_{\hat{y}}). \tag{4.7}$$

Considering that $\xi_{\hat{y}}$ and $\theta_{\hat{y}}$ are functions of $\Re(\hat{y})$, $\Im(\hat{y})$, $\Im(\hat{y}^*)$, and $|\Im(\hat{y})|$; for the sake of simplicity and compact representation, we next calculate these gradients before calculating the gradient of the cost function, $\nabla_{\mathbf{w}^{a*}} J$. The gradients of the real and imaginary parts of $\hat{y}$ are given by

$$\begin{aligned}
\nabla_{\mathbf{w}^{a*}} \Re(\hat{y}) =& \frac{1}{2}\nabla_{\mathbf{w}^{a*}} \left(\hat{y} + \hat{y}^*\right) = \frac{1}{2}\nabla_{\mathbf{w}^{a*}} \left(\mathbf{w}^{aH}\mathbf{q}^a + \mathbf{q}^{aH}\mathbf{w}^a\right) \\
=& \frac{1}{2}\nabla_{\mathbf{w}^{a*}} \left(\mathbf{w}^{aH}\mathbf{q}^a + \frac{\mathbf{q}^{aH}}{2}\left(\mathbf{w}^{a*^i} + \mathbf{w}^{a*^j} + \mathbf{w}^{a*^k} - \mathbf{w}^{a*}\right)\right) \\
=& \frac{1}{2}\mathbf{q}^a - \frac{1}{4}\mathbf{q}^{a*} \\
\nabla_{\mathbf{w}^{a*}} \Im(\hat{y}) =& \frac{1}{2}\nabla_{\mathbf{w}^{a*}} \left(\hat{y} - \hat{y}^*\right) = \frac{1}{2}\nabla_{\mathbf{w}^{a*}} \left(\mathbf{w}^{aH}\mathbf{q}^a - \mathbf{q}^{aH}\mathbf{w}^a\right) \\
=& \frac{1}{2}\nabla_{\mathbf{w}^{a*}} \left(\mathbf{w}^{aH}\mathbf{q}^a - \frac{\mathbf{q}^{aH}}{2}\left(\mathbf{w}^{a*^i} + \mathbf{w}^{a*^j} + \mathbf{w}^{a*^k} - \mathbf{w}^{a*}\right)\right) \\
=& \frac{1}{2}\mathbf{q}^a + \frac{1}{4}\mathbf{q}^{a*}
\end{aligned} \tag{4.8}$$

while, considering that $\Im(\hat{y}^*) = -\Im(\hat{y})$ yields

$$\nabla_{\mathbf{w}^{a*}} \Im(\hat{y}^*) = -\nabla_{\mathbf{w}^{a*}} \Im(\hat{y}) = -\frac{1}{2}\mathbf{q}^a - \frac{1}{4}\mathbf{q}^{a*}. \tag{4.9}$$

Now, through substituting $|\Im(\hat{y})| = \sqrt{\Im(\hat{y})\Im(\hat{y}^*)}$, the gradient of $|\Im(\hat{y})|$ becomes

$$\begin{aligned}
\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})| = \nabla_{\mathbf{w}^{a*}} \sqrt{\Im(\hat{y})\Im^*(\hat{y})} =& \frac{\big(\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})\big)\Im^*(\hat{y}) + \Im(\hat{y})\big(\nabla_{\mathbf{w}^{a*}} \Im^*(\hat{y})\big)}{2|\Im(\hat{y})|} \\
=& \frac{-\big(\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})\big)\Im(\hat{y}) - \Im(\hat{y})\big(\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})\big)}{2|\Im(\hat{y})|} \\
=& \frac{-\Re\left(\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})\right)\Im(\hat{y}) + \langle \nabla_{\mathbf{w}^{a*}} \Im(\hat{y}) \cdot \Im(\hat{y})\rangle}{|\Im(\hat{y})|}
\end{aligned} \tag{4.10}$$

where $\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})$ and $\nabla_{\mathbf{w}^{a*}} \Im^*(\hat{y})$ are given in (4.8) and (4.9), respectively.

Considering that $|\Im(\hat{y})|, \Re(\hat{y}) \in \mathbb{R}$, application of the $\mathbb{HR}$-calculus to compute the gradient of $\theta_{\hat{y}}$ gives

$$\nabla_{\mathbf{w}^{a*}} \theta_{\hat{y}} = \frac{1}{1 + \left(\frac{|\Im(\hat{y})|}{\Re(\hat{y})}\right)^2}\nabla_{\mathbf{w}^{a*}} \frac{|\Im(\hat{y})|}{\Re(\hat{y})} = \frac{\Re^2(\hat{y})}{|\hat{y}|^2}\left(\frac{\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|}{\Re(\hat{y})} + |\Im(\hat{y})|\nabla_{\mathbf{w}^{a*}} \frac{1}{\Re(\hat{y})}\right) \tag{4.11}$$

where

$$\nabla_{\mathbf{w}^{a*}} \frac{1}{\Re(\hat{y})} = \frac{-1}{\Re^2(\hat{y})} \nabla_{\mathbf{w}^{a*}} \Re(\hat{y})$$

and the terms $\nabla_{\mathbf{w}^{a*}} \Re(\hat{y})$ and $\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|$ are respectively given by (4.8) and (4.10). In the same fashion, the gradient of $\xi_{\hat{y}}$ is given by

$$\nabla_{\mathbf{w}^{a*}} \xi_{\hat{y}} = \frac{\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})}{|\Im(\hat{y})|} + \Im(\hat{y}) \left(\nabla_{\mathbf{w}^{a*}} \frac{1}{|\Im(\hat{y})|} \right) \tag{4.12}$$

where

$$\nabla_{\mathbf{w}^{a*}} \frac{1}{|\Im(\hat{y})|} = \frac{-1}{|\Im(\hat{y})|^2} \nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|$$

and the terms $\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})$ and $\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|$ are respectively given by (4.8) and (4.10).

Finally, the gradient of the cost function can be calculated by replacing (4.11) and (4.12) into (4.7), to obtain the closed-form expression

$$\begin{aligned} \nabla_{\mathbf{w}^{a*}} J = -\Bigg[& \left(\frac{\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})}{|\Im(\hat{y})|} - \Im(\hat{y}) \frac{\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|}{|\Im(\hat{y})|^2} \right) \theta_{\hat{y}} \\ & + \zeta_{\hat{y}} \frac{\Re^2(\hat{y})}{|\hat{y}|^2} \left(\frac{\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|}{\Re(\hat{y})} - |\Im(\hat{y})| \frac{\nabla_{\mathbf{w}^{a*}} \Re(\hat{y})}{\Re^2(\hat{y})} \right) \Bigg] \varepsilon \\ - \varepsilon \Bigg[& \left(\frac{\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})}{|\Im(\hat{y})|} - \Im(\hat{y}) \frac{\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|}{|\Im(\hat{y})|^2} \right) \theta_{\hat{y}} \\ & + \zeta_{\hat{y}} \frac{\Re^2(\hat{y})}{|\hat{y}|^2} \left(\frac{\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|}{\Re(\hat{y})} - |\Im(\hat{y})| \frac{\nabla_{\mathbf{w}^{a*}} \Re(\hat{y})}{\Re^2(\hat{y})} \right) \Bigg] \end{aligned} \tag{4.13}$$

where the terms $\nabla_{\mathbf{w}^{a*}} \Re(\hat{y})$ and $\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|$ and are given in (4.8) and (4.10).

Note that in cases where the system can be modeled in a linear fashion that is $\hat{y} = \mathbf{h}^H \mathbf{q}$, by replacing $\mathbf{w}^a$ with $\mathbf{h}$ and replacing $\mathbf{q}^a$ with $\mathbf{q}$ in (4.6) through to (4.13), the strictly linear QLMP can be obtained using the same procedure. In addition, if the objective is to track the bearings of an object in three-dimensions, the signal is most conveniently modeled as a pure quaternion resulting in $\theta = \pi/2$. Therefore, $\varepsilon_n = (\xi_{\hat{y}} - \xi_y)\,\pi/2$ which leads to the gradient of the cost being simplified into

$$\begin{aligned} \nabla_{\mathbf{w}^{a*}} J = & -\frac{\pi^2}{4} \left[\left(\frac{\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})}{|\Im(\hat{y})|} - \Im(\hat{y}) \frac{\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|}{|\Im(\hat{y})|^2} \right) \right] (\xi_{\hat{y}} - \xi_y) \\ & - \frac{\pi^2}{4} (\xi_{\hat{y}} - \xi_y) \left[\left(\frac{\nabla_{\mathbf{w}^{a*}} \Im(\hat{y})}{|\Im(\hat{y})|} - \Im(\hat{y}) \frac{\nabla_{\mathbf{w}^{a*}} |\Im(\hat{y})|}{|\Im(\hat{y})|^2} \right) \right] \end{aligned} \tag{4.14}$$

where $\pi^2/4$ is essentially an scaling factor and can be incorporated into the adaptation gain, resulting in a more computationally efficient algorithm without loss of performance.

4.4 Geometric interpretation

Consider the estimates of the phase of the reference signal y_n, before and after the weight vector update respectively given by $\hat{y}_{n|n} = \mathbf{w}_n^{aH}\mathbf{q}_n^a$ and $\hat{y}_{n|n+1} = \mathbf{w}_{n+1}^{aH}\mathbf{q}_n^a$. The phase estimation improvement achieved in this way can be quantified as

$$\Delta\hat{y}_n = \hat{y}_{n|n+1} - \hat{y}_{n|n} = (\mathbf{w}_{n+1}^a - \mathbf{w}_n^a)^H\mathbf{q}_n^a$$

while replacing $\mathbf{w}_{n+1}^a - \mathbf{w}_n^a = -\mu\nabla_{\mathbf{w}_n^{a*}}J_n = \mu(\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\varepsilon_n + \mu\varepsilon_n(\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)$ obtained from the expressions in (4.5) and (4.7) yields

$$\begin{aligned}
\Delta\hat{y}_n =& -\mu\left(\nabla_{\mathbf{w}_n^{a*}}J_n\right)^H\mathbf{q}_n^a = \mu\Big((\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\varepsilon_n + \varepsilon_n(\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\Big)^H\mathbf{q}_n^a \\
=&2\mu\Big(\Re(\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\Re(\varepsilon_n) + \Re(\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\Im(\varepsilon_n) \\
&+ \Im(\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\Re(\varepsilon_n) - \langle\Im(\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\cdot\Im(\varepsilon_n)\rangle\Big)^H\mathbf{q}_n^a \\
=&2\mu\Big((\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\varepsilon_n - (\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n)\times\varepsilon_n\Big)^H\mathbf{q}_n^a.
\end{aligned} \tag{4.15}$$

Furthermore, from substituting $\nabla_{\mathbf{w}_n^{a*}}\varepsilon_n = (\nabla_{\mathbf{w}_n^{a*}}\xi_{\hat{y}_n})\theta_{\hat{y}_n} + \xi_{\hat{y}_n}(\nabla_{\mathbf{w}_n^{a*}}\theta_{\hat{y}_n})$ into the expression in (4.15) we have

$$\begin{aligned}
\Delta\hat{y}_n =&2\mu\theta_{\hat{y}_n}\Big((\nabla_{\mathbf{w}_n^{a*}}\xi_{\hat{y}_n})\varepsilon_n - (\nabla_{\mathbf{w}_n^{a*}}\xi_{\hat{y}_n})\times\varepsilon_n\Big)^H\mathbf{q}_n^a \\
&+ 2\mu\Big(\xi_{\hat{y}_n}(\nabla_{\mathbf{w}_n^{a*}}\theta_{\hat{y}_n})\varepsilon_n - \big(\xi_{\hat{y}_n}(\nabla_{\mathbf{w}_n^{a*}}\theta_{\hat{y}_n})\big)\times\varepsilon_n\Big)^H\mathbf{q}_n^a.
\end{aligned} \tag{4.16}$$

In order to quantify the improvement in the estimate of the phase when using the estimate obtained after the weight vector has been updated, we shall split the term $\Delta\hat{y}_n$ as expressed in (4.16), into two parts, the angle-only term, given by

$$\Delta\theta_n = 2\mu\Big(\underbrace{\xi_{\hat{y}_n}(\nabla_{\mathbf{w}_n^{a*}}\theta_{\hat{y}_n})\varepsilon_n - \big(\xi_{\hat{y}_n}(\nabla_{\mathbf{w}_n^{a*}}\theta_{\hat{y}_n})\big)\times\varepsilon_n}_{\nabla_{\mathbf{w}_n^{a*}}J_n^\theta}\Big)^H\mathbf{q}_n^a \tag{4.17}$$

and the unit vector term, given by

$$\Delta\xi_n = 2\mu\Big(\underbrace{\theta_{\hat{y}_n}(\nabla_{\mathbf{w}_n^{a*}}\xi_{\hat{y}_n})\varepsilon_n - \theta_{\hat{y}_n}(\nabla_{\mathbf{w}_n^{a*}}\xi_{\hat{y}_n})\times\varepsilon_n}_{\nabla_{\mathbf{w}_n^{a*}}J_n^\xi}\Big)^H\mathbf{q}_n^a \tag{4.18}$$

where it can be shown that the terms $\nabla_{\mathbf{w}_n^{a*}}J_n^\theta$ and $\nabla_{\mathbf{w}_n^{a*}}J_n^\xi$ are the gradients of the cost function, J_n, if $\xi_{\hat{y}_n}$ and $\theta_{\hat{y}_n}$ were treated as constants. Therefore, the term in (4.17) reduces the error of the angle estimate while treating the unit vector as a constant, the geometric view of which is illustrated in Figure 4.1, whereas the term in (4.18) reduces the error of the unit vector estimate while treating the angle as a constant, the geometric view of which is illustrated in Figure 4.2. Together, these two terms form the basis for phase estimation in the quaternion domain.

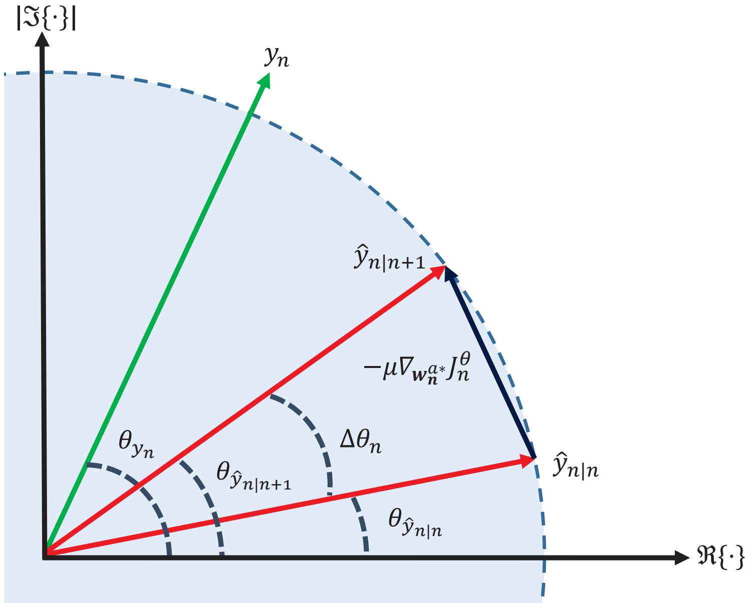

Figure 4.1: Geometric interpretation of the angle update in quaternion phase estimation. The desired signal, y_n, and its angle, θ_y, are shown relative to their phase-only estimates both before the weight vector update, $\{\hat{y}_{n|n}, \theta_{\hat{y}_{n|n}}\}$, and the after weight vector update, $\{\hat{y}_{n|n+1}, \theta_{\hat{y}_{n|n+1}}\}$. The direction of $-\mu\nabla_{\mathbf{w}_n^{a*}} J_n^{\theta}$ and $\Delta\theta_n$ as presented in (4.17) are also illustrated. The unit circle is shown in light blue as a visual reference.

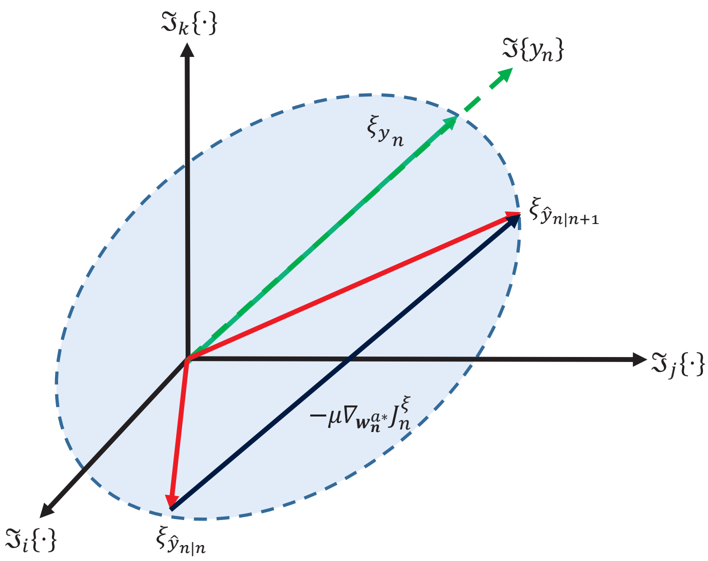

Figure 4.2: Geometric interpretation of the unit vector update in quaternion phase estimation. The desired signal, y_n, is shown relative to the unit vector of the phase-only estimate both before weight vector update, $\xi_{\hat{y}_{n|n}}$, and after weight vector updates, $\xi_{\hat{y}_{n|n+1}}$. Note that here $\Delta\xi_n = -\mu\nabla_{\mathbf{w}_n^{a*}} J_n^{\xi}$. A unit circle centered around the origin in the plane containing $\xi_{\hat{y}_{n|n}}$ and $\xi_{\hat{y}_{n|n+1}}$ is shown in light blue as a visual reference.

4.5 Stability analysis

The phase estimation errors before and after the weight vector update operation are respectively given by

$$\varepsilon_{n|n} = \xi_{\hat{y}_{n|n}} \theta_{\hat{y}_{n|n}} - \xi_y \theta_y$$
$$\varepsilon_{n|n+1} = \xi_{\hat{y}_{n|n+1}} \theta_{\hat{y}_{n|n+1}} - \xi_y \theta_y$$

the aim is to find the range for the adaption gain, μ, that ensures continuous learning that is $|\varepsilon_{n|n+1}|^2 < |\varepsilon_{n|n}|^2$. To this end, consider the The first-order Taylor expansion of $|\varepsilon_{n|n+1}|^2$ around $|\varepsilon_{n|n}|^2$ given by

$$\left|\varepsilon_{n|n+1}\right|^2 = \left|\varepsilon_{n|n}\right|^2 + \left\langle \nabla_{\mathbf{w}_n^{a*}} \left|\varepsilon_{n|n}\right|^2 \cdot \Delta \mathbf{w}_n^a \right\rangle \tag{4.19}$$

where $\Delta \mathbf{w}_n^a = \mathbf{w}_{n+1}^a - \mathbf{w}_n^a$. From the expression (4.5) we have $\Delta \mathbf{w}_n^a = -\mu \nabla_{\mathbf{w}_n^{a*}} J_n = -\mu \nabla_{\mathbf{w}_n^{a*}} \left|\varepsilon_{n|n}\right|^2$ that upon replacing into the expression in (4.19) yields

$$\begin{aligned} \left|\varepsilon_{n|n+1}\right|^2 &= \left|\varepsilon_{n|n}\right|^2 - \mu \left\langle \nabla_{\mathbf{w}_n^{a*}} \left|\varepsilon_{n|n}\right|^2 \cdot \nabla_{\mathbf{w}_n^{a*}} \left|\varepsilon_{n|n}\right|^2 \right\rangle \\ &= \left|\varepsilon_{n|n}\right|^2 - \mu \left| \nabla_{\mathbf{w}_n^{a*}} \left|\varepsilon_{n|n}\right|^2 \right|^2 . \end{aligned} \tag{4.20}$$

Through substituting the expression in (4.6) into (4.20) we have

$$\begin{aligned} \left|\varepsilon_{n|n+1}\right|^2 &= \left|\varepsilon_{n|n}\right|^2 - \mu \left| (\nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n}) \varepsilon_{n|n} + \varepsilon_{n|n} (\nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n}) \right|^2 \\ &= \left|\varepsilon_{n|n}\right|^2 - 4\mu \left| \varepsilon_{n|n} \Re(\nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n}) - \left\langle \varepsilon_{n|n} \cdot \Im(\nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n}) \right\rangle \right|^2 \end{aligned}$$

where considering that $\left\langle \varepsilon_{n|n} \cdot \Im(\nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n}) \right\rangle \in \mathbb{R}$ and that $\varepsilon_{n|n} \Re(\nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n})$ lies solely in the quaternion imaginary subspace gives

$$\begin{aligned} \left|\varepsilon_{n|n+1}\right|^2 &= \left|\varepsilon_{n|n}\right|^2 - 4\mu \left|\varepsilon_{n|n}\right|^2 \left| \Re(\nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n}) \right|^2 - 4\mu \left|\varepsilon_{n|n}\right|^2 \left| \Im(\nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n}) \right|^2 \\ &= \left|\varepsilon_{n|n}\right|^2 \left(1 - \mu 4 \left| \nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n} \right|^2 \right) . \end{aligned}$$

Therefore, in order to ensure contentious learning of the developed quaternion phase-only estimator, $|\varepsilon_{n|n+1}|^2 < |\varepsilon_{n|n}|^2$, it suffices to grantee that

$$\left| 1 - 4\mu \left| \nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n} \right|^2 \right| < 1$$

resulting in the following bound for the adaptation gain

$$0 < \mu < \frac{1}{2 \left| \nabla_{\mathbf{w}^{a*}} \varepsilon_{n|n} \right|^2} .$$

4.6 Simulations

In this section, the performance of the developed quaternion phase estimator is validated through simulations in practical applications. The QLMP is implemented for tracking rotations of an object in the three-dimensional domain and for estimating the main frequency of three-phase power systems. In addition, the WL-QLMP is implemented for tracking the limbs of a person performing Tai-Chi movements.

4.6.1 Synthetically generated three-dimensional data

In this section we consider unit amplitude three-dimensional signals modeled as pure quaternions. The objective is to track the rotations that the signal undergoes from the point of view of an observer at the origin, a classic problem encountered in attitude estimation and bearings-only tracking. In this scenario, since the observer is at the center of the coordinate system, the rotation vector, $\boldsymbol{\eta}$, will always be normal to plane of motion that encompasses the center of the coordinate system, the pre-rotation position of the object, and the post-rotation position of the object. Thus, $\langle q_{\boldsymbol{\eta}} \cdot q_{\mathbf{u}} \rangle = \langle q_{\boldsymbol{\eta}} \cdot q_{\mathbf{u}'} \rangle = 0$, and the quaternion rotation model expressed in (4.1) yields

$$
\begin{aligned}
q_{\mathbf{u}'} &= \left(\cos\left(\frac{\theta}{2}\right) + q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\right) q_{\mathbf{u}} \left(\cos\left(\frac{\theta}{2}\right) - q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\right) \\
&=\cos^2\left(\frac{\theta}{2}\right) q_{\mathbf{u}} + 2q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\cos\left(\frac{\theta}{2}\right) q_{\mathbf{u}} - \sin^2\left(\frac{\theta}{2}\right) q_{\mathbf{u}} \\
&= \left(\cos\left(\frac{\theta}{2}\right) + q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\right)^2 q_{\mathbf{u}}.
\end{aligned}
\tag{4.21}
$$

Therefore, considering the expression in (4.21), in these scenarios, the rotation can be tracked using the QLMP where the filter coefficient is given by

$$
h = \left(\cos\left(\frac{\theta}{2}\right) + q_{\boldsymbol{\eta}}\sin\left(\frac{\theta}{2}\right)\right)^2 = e^{q_{\boldsymbol{\eta}}\theta}
$$

from which the axis and angle of rotation can easily be calculated and are given by

$$
q_{\boldsymbol{\eta}} = \frac{\Im\left(\ln(h)\right)}{\left|\Im\left(\ln(h)\right)\right|} \qquad \text{and} \qquad \theta = \left|\Im\left(\ln(h)\right)\right|. \tag{4.22}
$$

In the first simulation, the performance of the developed quaternion phase only estimator was validated using synthetically generated signals with unit amplitude. The signals were set to oscillate in a randomly selected plane of the quaternion imaginary subspace with frequency of 20 Hz. The signal was considered to be corrupted by additive white Gaussian noise (AWGN) with signal to noise ratio (SNR) of 35 dB and the sampling frequency was considered to be 1 KHz. The QLMP, QLMS, and a real-valued quadrivariate LMS were implemented in order to track the signal. This experiment was repeated 100 times. The MSE performance of the QLMP is shown alongside the MSE performance of the QLMS and the real-valued quadrivariate LMS in Figure 4.3. In addition, the signal from one realization of the experiment and its estimates obtained through implementing the QLMP, QLMS, and real-valued quadrivariate LMS is shown in Figure 4.4. Observe that the developed QLMP algorithm not only converged faster, but also achieved a lower steady-state MSE.

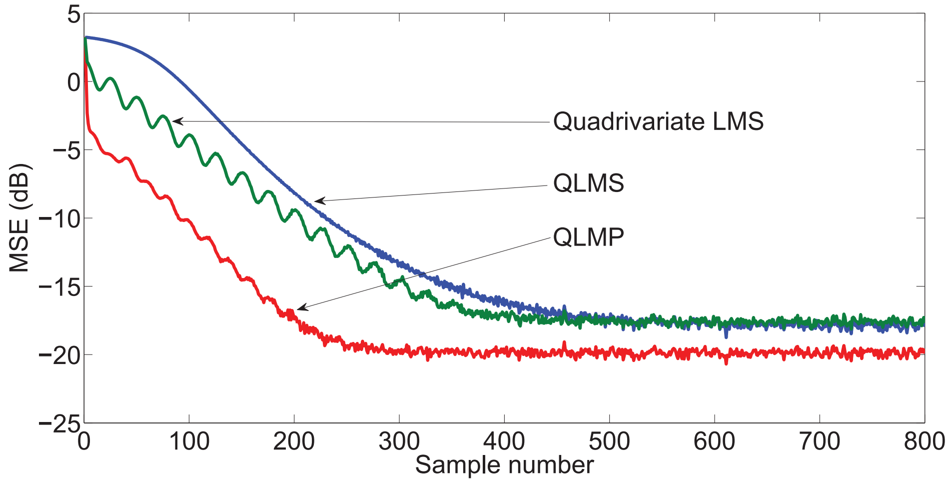

Figure 4.3: MSE performance of the developed QLMP algorithm compared to that of the QLMS and real-valued quadrivariate LMS, when tracking an object oscillating in a randomly selected plane of the quaternion imaginary subspace.

In the second simulation, the tracking ability of the QLMP algorithm is demonstrated, where the QLMP algorithm was implemented to track the changing axis and angle of rotation of an object rotating around the center of the coordinate system. Figure 4.5 shows the instantaneous rotation angle and its estimate, whereas the rotation vector and its estimate are shown in Figure 4.6. Note that the QLMP accurately tracked both the rotation vector and rotation angle.

4.6.2 Power system frequency estimation

Since their introduction, analytical signals have been an integral part of spectral analysis of real-valued signals. In essence, analytical signals are complex-valued signals with one-sided Fourier transforms, generated by adding a real-valued signal to its Hilbert transform, where the phase contains the spectral information of the real-valued signal. However, in circumstances where the signal of interest is complex-valued employing the Hilbert transform in the complex domain, due to the limited dimensionality of complex numbers, will lead to the loss of the information contained in the negative part of the spectrum. Therefore, in order to fully characterize complex-valued signals, the concept of analytical signal has been expanded to the quaternion domain in [72, 73] to present a

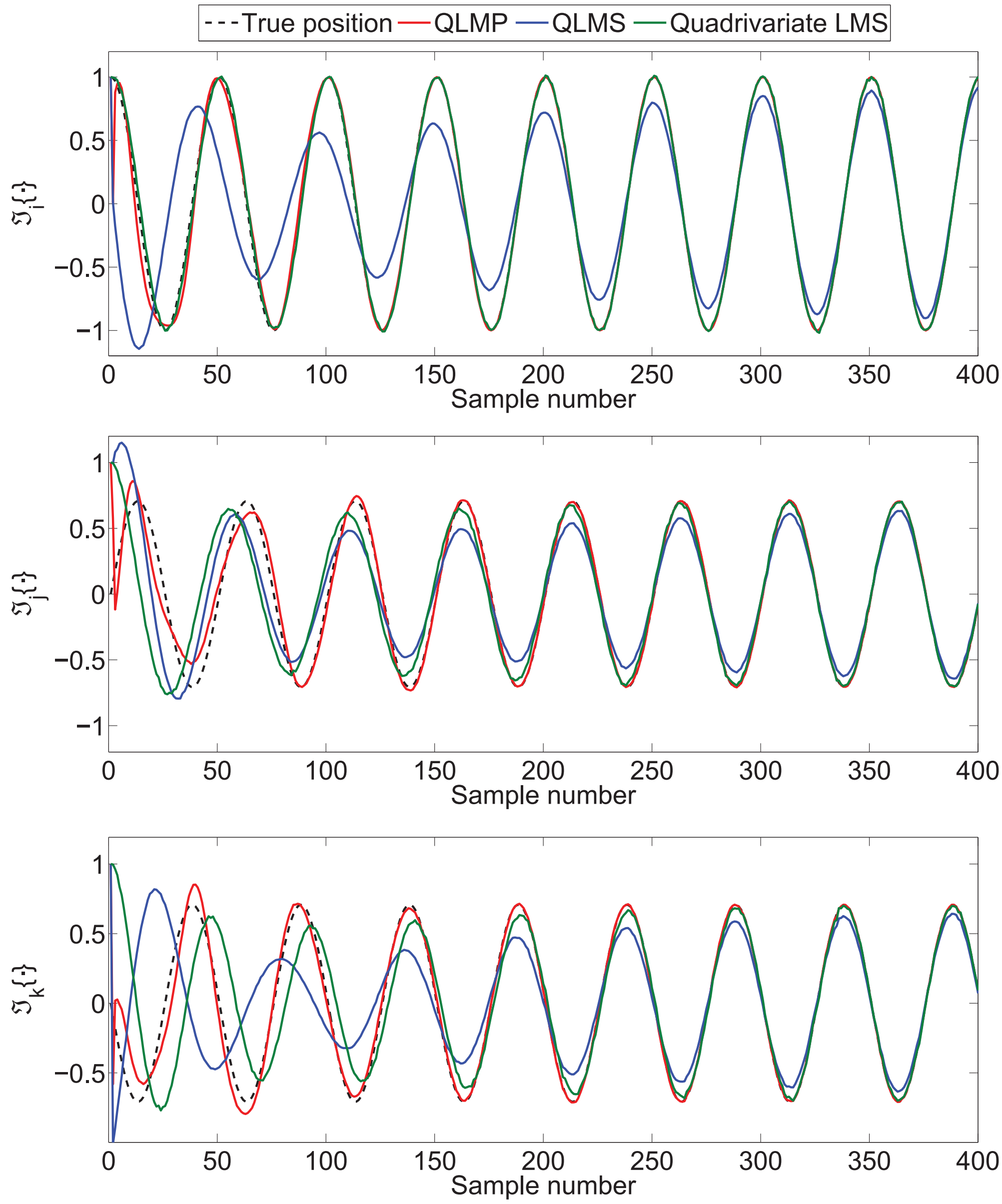

Figure 4.4: Tracking a pure imaginary quaternion-valued signal with unit amplitude through implementing the developed QLMP algorithm (in red), the QLMS algorithm (in blue), and the real-valued quadrivariate LMS (in dark green).

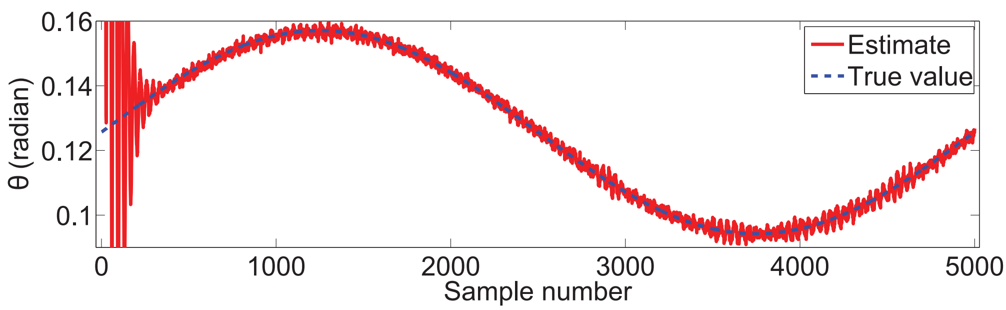

Figure 4.5: Rotation angle estimation of a pure imaginary quaternion-valued signal with unit amplitude through implementing the QLMP algorithm.

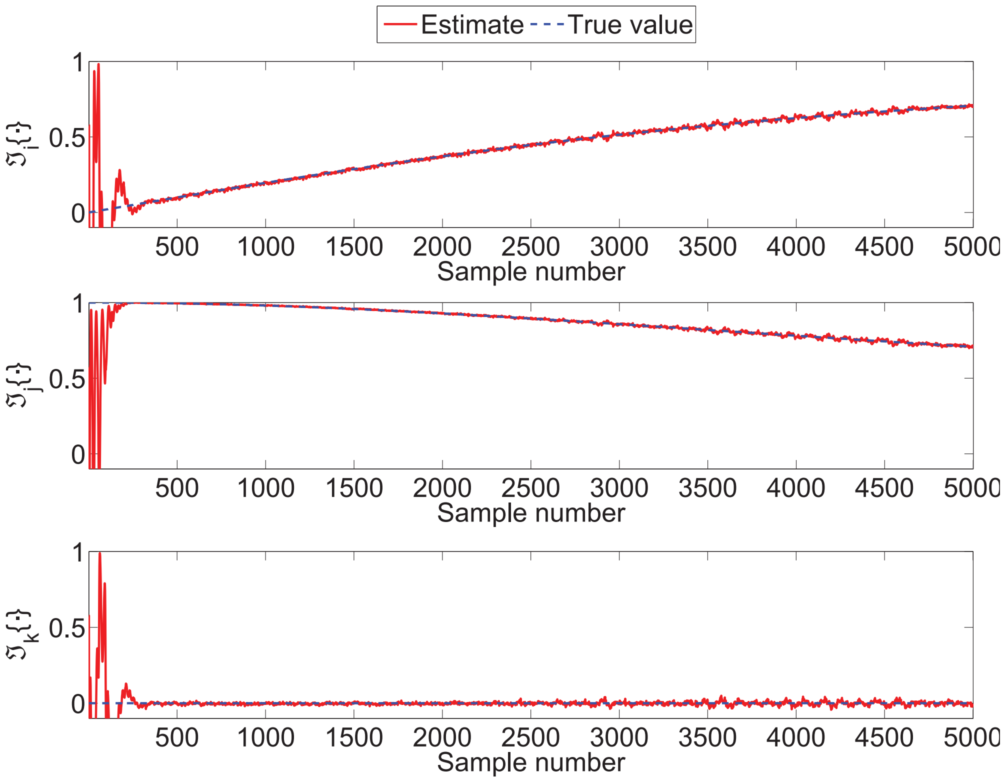

Figure 4.6: Estimates of rotation vector, a pure imaginary quaternion-valued signal with unit amplitude, obtained through implementing the QLMP algorithm.

quaternion-valued representation for complex-valued signals with one-sided quaternion Fourier transforms and it has been shown that the spectral information of the complex-valued signal is contained in the phase of its quaternion-valued representation. Based on these developments, we next present a simple method for estimating the main frequency component of three-phase power system signals.

For spectral analysis of three-phase power systems, the Clarke transform is used to map the three-phase voltages onto the complex domain, where in general they are shown to trace an ellipse [56]. The Hilbert transform of the output of the Clarke transform, v_n, with respect to the imaginary unit $\zeta \in \{i, j, k\}$ given by $\mathcal{H}_\zeta(v_n)$, is used to construct the phase-only signal

$$S_n = \frac{v_n + \mathcal{H}_\zeta(v_n)}{|v_n + \mathcal{H}_\zeta(v_n)|}.$$

The instantaneous frequency of the system is now given by

$$f = \frac{1}{2\pi\zeta}\ln\left(\frac{\hat{S}_{n+1}}{S_n}\right)f_s$$

where f_s represents the sampling frequency, and $\hat{S}_{n+1}$ is the estimate of S_{n+1} obtained using the strictly linear QLMP (for details on implementation of the Hilbert transform see Appendix B).

In the first experiment, the proposed algorithm was used to estimate the fundamental frequency, $f = 50$ Hz, of a three-phase power system where $f_s = 1500$ Hz and it was assumed that the measurements were corrupted by AWGN with SNR of 50 dB. The system was operating under balanced (nominal) conditions for the first 2 seconds, then it was forced into unbalanced operating condition by an 80% reduction in the amplitude of one of its phases, which continued for 2 seconds, and for the last 2 seconds one of the phases of the three-phase power systems experienced a 1 Hz amplitude modulation given by $1 + 0.2\sin(2\pi\Delta T n)$. Figure 4.7 shows the system frequency and its estimate along side the frequency estimation error performance of the proposed method given by

$$\varepsilon_n = f_n - \hat{f}_n$$

where f_n and $\hat{f}_n$ represent the system frequency and its estimate at time instant n. Note that the proposed method accurately tracked the frequency of the power system under both balanced and unbalanced conditions. In addition, unbalanced operating conditions and amplitude modulation did not significantly effect the frequency estimation error performance of the proposed method.

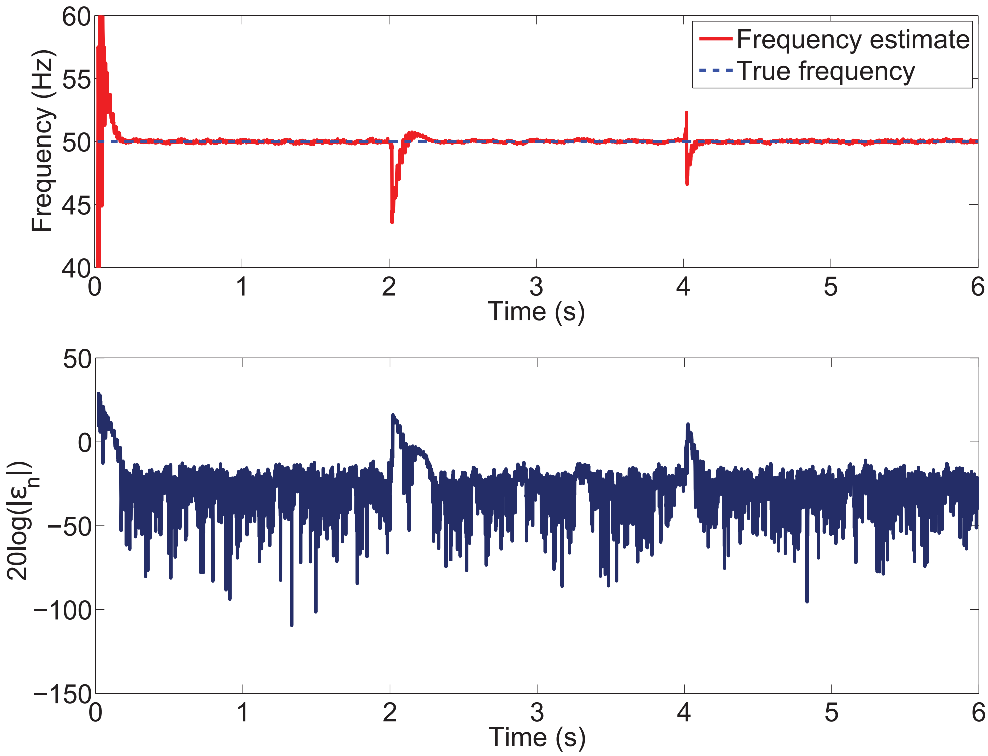

Figure 4.7: Frequency estimation for balanced and unbalanced three-phase power system; frequency and its estimation (top), squared error $|\varepsilon_n|^2 = \left|f_n - \hat{f}_n\right|^2$ (bottom). The system is balanced for the first two seconds, is unbalanced for the next two second seconds, and experiences an amplitude modulation on one of its phases for the last two seconds.

The performance of the proposed algorithm was also investigated for the case where the power system has a rising (*cf.* falling) fundamental frequency, a typical case where power generation is higher (*cf.* lower) than consumption, while operating under unbalanced condition caused by an 80% reduction in the amplitude of one phase. Figure 4.8 shows the system frequency and its estimate along side the frequency estimation error performance of the proposed method, verifying that the proposed algorithm can accurately estimate the fundamental frequency of the system. In addition, it is important to observe that form Figure 4.8, it becomes apparent that the developed QLMP algorithm can track the phase information of quaternion-valued signals even in incidences when the phase is experiencing rapid changes. Finally, note that in three-phase power systems frequency changes at much lower rates than is considered in this simulation and the rates selected are to show the performance of the algorithm under extreme or worst case scenarios.

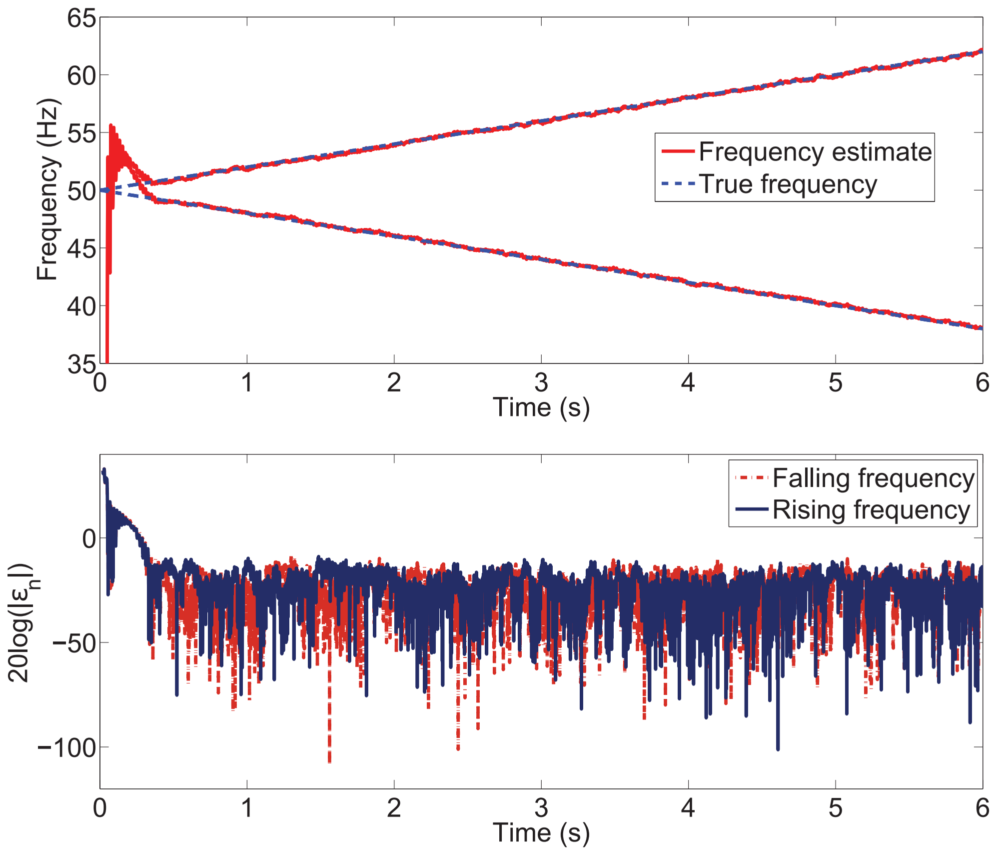

Figure 4.8: Frequency estimation for a power system experiencing frequency change at the rate of 2 Hz/s; frequency and its estimate (top), squared error $|\varepsilon_n|^2 = \left|f_n - \hat{f}_n\right|^2$ (bottom).

4.6.3 Body-motion tracking

The WL-QLMP was next used in a one-step-ahead prediction setting in order to track the rotations of limbs of an athlete preforming Tai-Chi movements. The data were recorded using four accelerometers, which measured the three Euler angles α, β, and γ, as shown in Figure 4.9, these angles respectively represent the roll, pitch, and yaw within the range $[\pi, -\pi]$. The Euler angles can be transformed into rotation matrices; however, these matrices contain singularities when one of the Euler angles approaches $\pm\frac{\pi}{2}$, the phenomenon known as the gimbal lock.

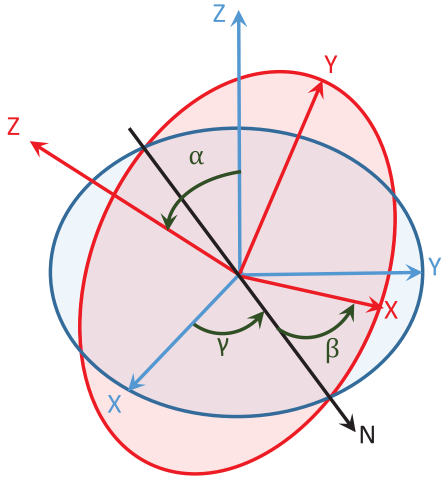

Figure 4.9: Inertial body motion sensor setting: fixed coordinate system (blue), sensor coordinate system (red), and Euler angles (green). The "N" axis is used as a visual guide to indicate the yaw angle.

In order to make the data suitable for processing[1] each angle was converted into a quaternion using the following transformations

$$\begin{aligned} \alpha &\to \big(\cos(\alpha), \sin(\alpha)\big) \to e^{i\frac{\alpha}{2}} \\ \beta &\to \big(\cos(\beta), \sin(\beta)\big) \to e^{j\frac{\beta}{2}} \\ \gamma &\to \big(\cos(\gamma), \sin(\gamma)\big) \to e^{k\frac{\gamma}{2}} \end{aligned} \tag{4.23}$$

where each term represents the roll, pitch, and yaw rotations in the three-dimensional space respectively. The three quaternion terms obtained in (4.23) were combined into one quaternion given by

$$q = \big(e^{k\frac{\gamma}{2}}\big)\big(e^{j\frac{\beta}{2}}\big)\big(e^{i\frac{\alpha}{2}}\big) \tag{4.24}$$

which represents the total rotation. As the mapping in (4.23) is invertible, the resulting quaternion in (4.24) preserves the dynamics of the recorded signal. Figure 4.10 shows the recorded yaw angle measurements containing discontinuities and its corresponding continuous transformation in (4.23) which is ready for processing.

[1]The recorded angles are unsuitable for processing in their original form as they contain discontinuities around $\pm\pi$.

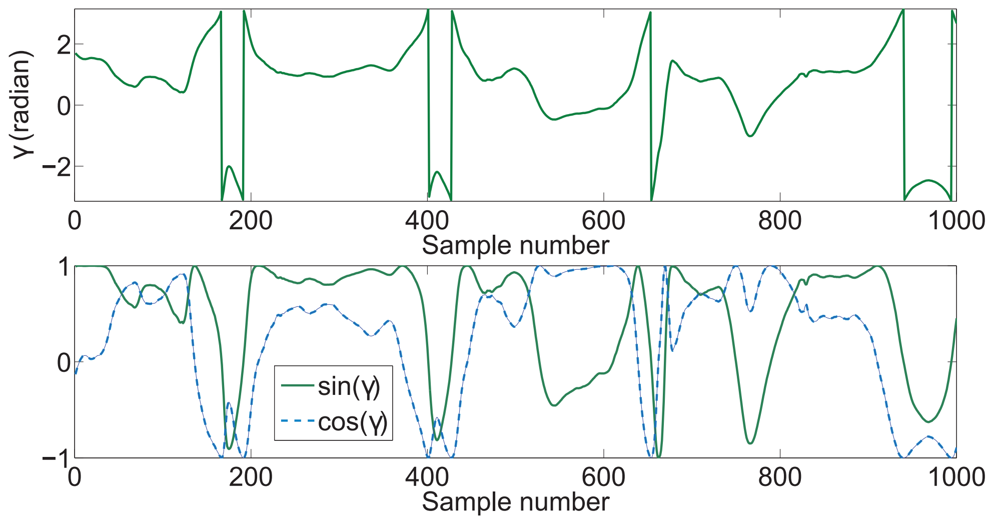

Figure 4.10: Recorded yaw angle measurements with discontinuities at $\pm\pi$ (top) and its contentious transform (bottom).

The WL-QLMP algorithm was considered in a one-step-ahead prediction setting to track body rotations. In Figure 4.11, the estimates of the absolute value of the phase of the body motion signal obtained by the WL-QLMP algorithm are shown alongside those obtained through a real-valued quadrivariate approach, where the quaternion-valued body motion signal was considered in its vector representation and the real-valued LMS algorithm was employed to track the signal. Notice that the WL-QLMP algorithm has a significantly better performance than that of the quadrivariate algorithm. The individual components of the phase of the quaternion-valued body motion signal are shown in Figure 4.12, which confers that the proposed WL-QLMP can be used to accurately track the rotations of an object moving in three-dimensions. In addition, from Figure 4.11 and Figure 4.12 observe that the WL-QLMP and the quadrivariate algorithm implemented to track the body motion signal seem to lose track of the signal during the terminating samples. This is mostly due to the sudden change in the body motion signal statistics or in other words the person performing the Tai-Chi movements coming to a standstill at the end of experiment.

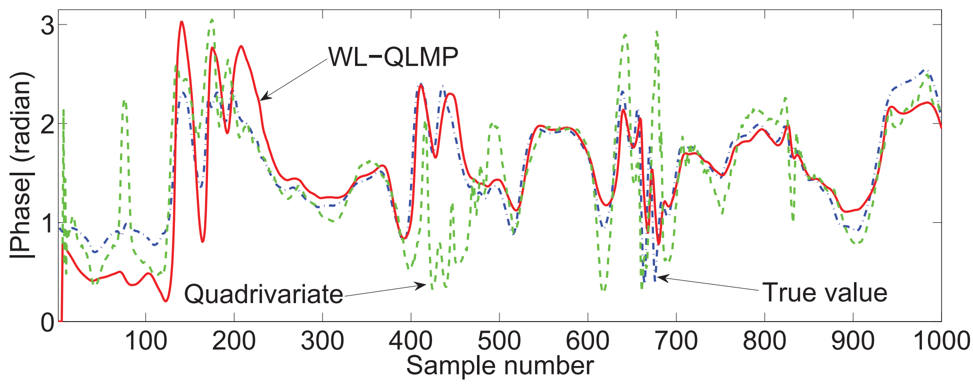

Figure 4.11: Absolute value of the phase of quaternion-valued body motion signal employing the WL-QLMP algorithm and a quadrivariate approach.

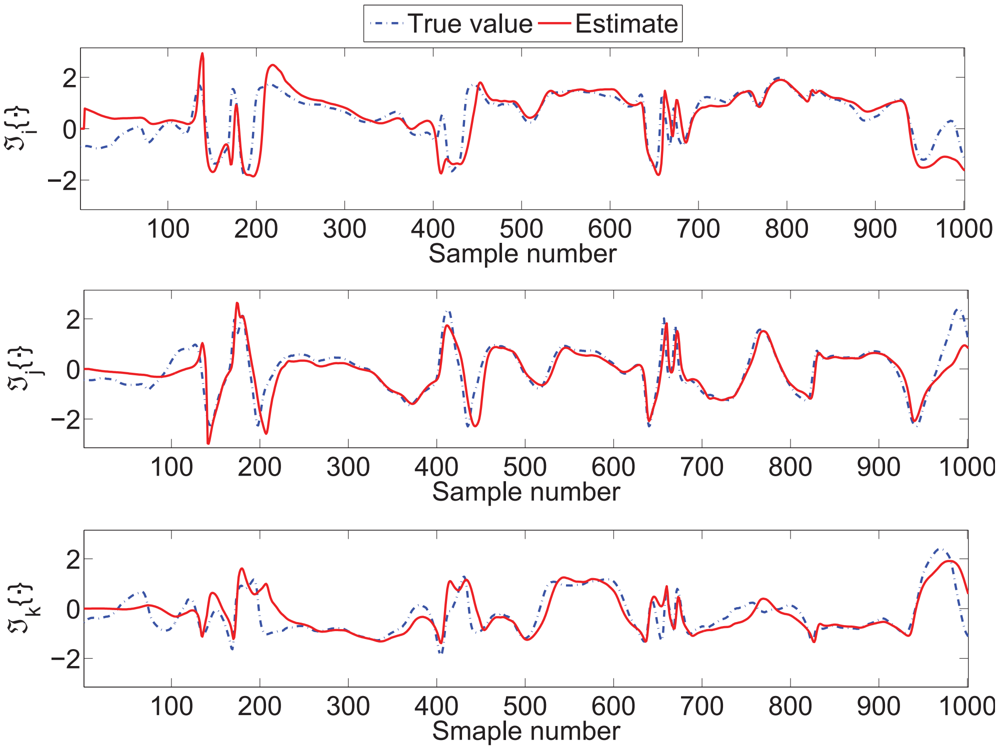

Figure 4.12: Phase estimation for recorded body-motion signal employing the WL-QLMP algorithm.

4.7 Conclusion

The WL-QLMP adaptive filter has been introduced for the unified processing of three and four-dimensional phase-only signals. The WL-QLMP has been shown to account for both second-order proper and improper four-dimensional data represented as full quaternions. The performance of the algorithms has been analyzed, a geometrical interpretation of the operations of the proposed algorithms has been presented, and convergence conditions have been established. Furthermore, the proposed algorithm has been validated in a number of practical applications including frequency estimation in three-phase power systems and body-motion tracking.

Chapter 5

A Non-Linear Complex-Valued Frequency Estimator for Three-Phase Power Systems

5.1 Overview

In this chapter, frequency estimation in three-phase power systems is considered from a state space point of view in order to develop a robust and fast converging algorithm for estimating the fundamental frequency of three-phase power systems in real-time. To this end, the Clarke transform is used to incorporate the information from all the phases into a complex-valued signal; then, a complex-valued widely-linear state space estimator that can accurately estimate the fundamental frequency of both balanced and unbalanced three-phase power systems is designed. Furthermore, it is shown that the framework can be extended to account for harmonic contaminations in the system. For rigor, the performance of the developed frequency estimator is quantified and compared to that of its counterparts. The performance of the developed algorithm is validated through simulations on both synthetic data and real-world data recordings, where it is shown that the developed algorithm outperforms both standard linear and the recently introduced widely-liner complex-valued frequency estimators.

5.2 Introduction

The components of the power grid are designed to operate optimally at a given nominal frequency [106]. Large deviations from the nominal system frequency adversely affects the components of the power grid [106–108], such as compensators and loads, resulting in harmful operating conditions that can propagate throughout the network and cause stability issues. Thus, making frequency stability one of the most important factors in power quality [106]. Therefore, accurate frequency estimation is a prerequisite to ensuring frequency stability of the grid and maintaining optimal operating conditions.

The importance of frequency estimation in power grids has motivated the introduction of a variety of algorithms for this purpose, including frequency estimation techniques based on the use of phase-locked loops [109, 110], recursive Newton-type frequency and spectrum estimation algorithms applicable to three-phase power systems [111, 112], approaches based on the least squares and least mean square algorithms [113, 114], a Fourier transform based method for estimating the main frequency in three-phase power systems [115], and an adaptive notch filter for direct estimation of frequency and its rate of change in three-phase power systems [116]. In particular, approaches based on the Kalman and extended Kalman filters have been shown to be advantageous [117–119], due to their ability to model observational noise. Although a great deal of research has been conducted in this area, there are shortcomings that are summarized in the following:

- Frequency estimation techniques based on phase-locked loops and notch filters are computationally intensive; in addition, in the case of phase-locked loops, dedicated hardware is also required.
- Frequency estimation techniques that use block based estimators, such as the least squares and the Fourier transform, are not adequate for signals with rapidly changing statistics such as those encountered in power systems that are starting to experience a fault or are recovering from one.
- In many incidences [111, 113, 117, 118], the frequency estimator only uses voltage measurements from a single phase and cannot fully characterize three-phase systems, especially during crucial moments when one or two of the phases encounter a sudden drop in voltage or short circuit.
- Frequency estimators based on standard complex-valued linear models [114, 119] are shown to experience large oscillatory errors at twice the frequency of the system when the three-phase system is operating in an unbalanced fashion [116, 120].

In order to introduce a robust frequency estimator for three-phase power systems, the Clarke transform and widely-linear modeling of complex-valued signals have been used in [57, 121], where an algorithm based on the ACLMS adaptive filter has been presented. In addition, the model has been employed in its state space formulation in [58, 122] to introduce a frequency estimator based on the AECKF filter that outperforms its ACLMS based counterpart, due to the fact that it can account for observational noise. Although these algorithms perform significantly better than their strictly linear counterparts, their performance deteriorates when the power system is operating under fault conditions.

In this chapter, frequency estimation in three-phase power systems is considered from a new perspective. The analysis shows that the output of the Clarke transform of a balanced three-phase system comprises a positive sequence only; however, when the three-phase system is operating under fault conditions the output of the Clarke transform comprises both a positive and a negative sequence element with the same fundamental frequency. This is used to design a widely-linear state space model that accounts for the presence of the negative sequence and can operate optimally under both balanced and unbalanced operating conditions. The resulting frequency estimator is computationally efficient, unbiased, and has consistent performance regardless of operating conditions. Furthermore, we show that the framework can be extended to account for the presence of the main harmonic components of the power signal.

5.3 Three-phase power systems

The instantaneous voltages of each phase in a three-phase power system are given by [53]

$$\begin{aligned} v_{a,n} =& V_{a,n}\cos(2\pi f \Delta T n + \phi_{a,n}) \\ v_{b,n} =& V_{b,n}\cos\big(2\pi f \Delta T n + \phi_{b,n} + \frac{2\pi}{3}\big) \\ v_{c,n} =& V_{c,n}\cos\big(2\pi f \Delta T n + \phi_{c,n} + \frac{4\pi}{3}\big) \end{aligned} \tag{5.1}$$

where $V_{a,n}$, $V_{b,n}$, and $V_{c,n}$ are instantaneous amplitudes, $\phi_{a,n}$, $\phi_{b,n}$, and $\phi_{c,n}$ are instantaneous phases, f is the system frequency, and $\Delta T = 1/f_s$ is the sampling interval with f_s denoting the sampling frequency. The Clarke transform, given by [53]

$$\begin{bmatrix} v_{0,n} \\ v_{\alpha,n} \\ v_{\beta,n} \end{bmatrix} = \sqrt{\frac{2}{3}} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 1 & -\frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} v_{a,n} \\ v_{b,n} \\ v_{c,n} \end{bmatrix} \tag{5.2}$$

maps the three-phase power system onto a new domain where they are represented by $v_n = v_{\alpha,n} + jv_{\beta,n}$ while in most practical applications $v_{0,n}$ is ignored and only serves the role of making the Clarke transform reversible. When the three-phase system is operating under blanched conditions that is $V_n = V_{a,n} = V_{b,n} = V_{c,n}$ and $\phi_{a,n} = \phi_{b,n} = \phi_{c,n} = \phi_n$, it is straightforward to show $v_{0,n} = 0$, which in turn results in

$$v_n = \sqrt{\frac{3}{2}} V_n e^{i(2\pi f \Delta T n + \phi_n)} \tag{5.3}$$

that can be expressed by employing the first order linear autoregressive model

$$v_n = e^{i2\pi f \Delta T} v_{n-1}$$

where the term $e^{i2\pi f \Delta T}$ is referred to as the phase incrementing element.

The expression in (5.3) shows that when the three-phase power system is balanced, v_n is consisted of only a positive sequenced element; hence, it will trace a circle on the complex plane making the distribution of v_n rotation invariant or proper. Moreover, under balanced operating condition the frequency of the system can be estimated by standard linear complex-valued Kalman filters employing the state space model given in Algorithm 4, where $\varphi_n = e^{i2\pi f \Delta T}$ is the phase incrementing element [119].

Algorithm 4. Complex Linear Frequency Estimator (CLFE) [119]

State evolution equation:

$$\begin{bmatrix} \varphi_n \\ \hat{v}_n \end{bmatrix} = \begin{bmatrix} \varphi_{n-1} \\ \varphi_{n-1}\hat{v}_{n-1} \end{bmatrix} + \boldsymbol{\nu}_n$$

Observation equation:

$$v_n = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \varphi_n \\ \hat{v}_n \end{bmatrix} + \omega_n$$

Estimate of frequency:

$$\hat{f}_n = \frac{\Im\left(\ln\left(\varphi_n\right)\right)}{i2\pi\Delta T}$$

The power system is designed to operate optimally at its nominal frequency and under balanced conditions; nonetheless, in practice, a wide range of phenomena, such as voltage sags, load imbalance, and faults in the transmission line, will lead to unbalanced operating conditions in three-phase power systems [107, 108]. When experiencing unbalanced operating conditions [121]

$$v_n = A_n e^{i(2\pi f \Delta T n + \phi_n)} + B_n e^{-i(2\pi f \Delta T n + \phi_n)}$$

with

$$\begin{aligned} A_n =& \frac{\sqrt{6}\left(V_{a,n} + V_{b,n} + V_{c,n}\right)}{6} \\ B_n =& \frac{\sqrt{6}\left(2V_{a,n} - V_{b,n} - V_{c,n}\right)}{12} - i\frac{\sqrt{2}\left(V_{b,n} - V_{c,n}\right)}{4}. \end{aligned}$$

where the authors in [121] have considered all phase shifts to be equal to ϕ_n for the sake of simplicity. Therefore, v_n comprises both a positive and a negative sequenced element and will trace an ellipse in the complex plane, making the distribution of v_n improper.

In order to accommodate both balanced and unbalanced systems, it has been shown that v_n can be expressed by employing the first order widely-linear autoregressive model

$$v_n = h_{n-1}v_{n-1} + g_{n-1}v^*_{n-1}$$

where h_n and g_n are the linear and conjugate coefficients respectively [121]. The fundamental frequency of both balanced and unbalanced three-phase power systems can now be estimated by a AECKF employing the state space model given in Algorithm 5 [122].

Algorithm 5. Complex Widely-Linear Frequency Estimator (CWLFE) [122]

State evolution equation:

$$\begin{bmatrix} h_n \\ g_n \\ \hat{v}_n \\ h_n^* \\ g_n^* \\ \hat{v}_n^* \end{bmatrix} = \begin{bmatrix} h_{n-1} \\ g_{n-1} \\ h_{n-1}\hat{v}_{n-1} + g_{n-1}\hat{v}_{n-1}^* \\ h_{n-1}^* \\ g_{n-1}^* \\ h_{n-1}^*\hat{v}_{n-1}^* + g_{n-1}^*\hat{v}_{n-1} \end{bmatrix} + \boldsymbol{\nu}_n^a$$

Observation equation:

$$\begin{bmatrix} v_n \\ v_n^* \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_n \\ g_n \\ \hat{v}_n \\ h_n^* \\ g_n^* \\ \hat{v}_n^* \end{bmatrix} + \boldsymbol{\omega}_n^a$$

Estimate of frequency:

$$\hat{f}_n = \frac{\arcsin\left(-i\Im\left(h_n + a_n\right)\right)}{2\pi\Delta T}$$

with

$$a_n = -\Im\left(h_n\right) + \sqrt{\Im^2\left(h_n\right) + |g_n|^2}$$

Note that for a balanced three-phase system $h_n = e^{i2\pi f\Delta T}$ and $g_n = 0$; thus, under balanced operating conditions Algorithm 5 and Algorithm 4 will essentially operate akin to each other with the exception that Algorithm 4 cannot account for the presence of improper observational noise, due to its underlying linear model. However, observe that in Algorithm 5 the system frequency is calculated as a function of the filter coefficients, h_n and g_n, which not only increases the computational complexity of the algorithm, but also has a detrimental effect on its performance especially when the system is operating under unbalanced conditions. To address these issues, a novel complex-valued state space model for frequency estimation in three-phase power systems is presented that has consistent performance under both balanced and unbalanced operating conditions.

5.4 Frequency estimation using the positive and negative sequence elements

From the expression in (5.1) and (5.2), after some tedious mathematical manipulations, the output of the Clarke transform for a general three-phase system can be expressed as

$$v_n = \Lambda_{I,n}\cos(2\pi f \Delta T n) - \Lambda_{Q,n}\sin(2\pi f \Delta T n)$$

where $\{\Lambda_{I,n}, \Lambda_{Q,n}\} \in \mathbb{C}$ and are given by

$$\begin{aligned}
\Lambda_{I,n} =& \sqrt{\frac{2}{3}}V_{a,n}\cos(\phi_{a,n}) + (\frac{i\sqrt{3}-1}{\sqrt{6}})V_{b,n}\cos\big(\phi_{b,n} + \frac{2\pi}{3}\big) \\
& - (\frac{i\sqrt{3}+1}{\sqrt{6}})V_{c,n}\cos\big(\phi_{c,n} + \frac{4\pi}{3}\big) \\
\Lambda_{Q,n} =& \sqrt{\frac{2}{3}}V_{a,n}\sin(\phi_{a,n}) + (\frac{i\sqrt{3}-1}{\sqrt{6}})V_{b,n}\sin\big(\phi_{b,n} + \frac{2\pi}{3}\big) \\
& - (\frac{i\sqrt{3}+1}{\sqrt{6}})V_{c,n}\sin\big(\phi_{c,n} + \frac{4\pi}{3}\big).
\end{aligned}$$

Now, substituting the $\sin(\cdot)$ and $\cos(\cdot)$ with their polar representations yields

$$v_n = \underbrace{(\frac{\Lambda_{I,n}}{2} - \frac{\Lambda_{Q,n}}{2i})e^{i2\pi f\Delta T n}}_{v_n^+} + \underbrace{(\frac{\Lambda_{I,n}}{2} + \frac{\Lambda_{Q,n}}{2i})e^{-i2\pi f\Delta T n}}_{v_n^-}$$

where v_n has been decomposed into two counter rotating elements, v_n^+ with only a positive and v_n^- with only a negative sequenced element. The two counter-rotating elements can be modeled individually by employing the linear recursive models

$$v_n^+ = e^{i2\pi f\Delta T}v_{n-1}^+ \qquad \text{and} \qquad v_n^- = e^{-i2\pi f\Delta T}v_{n-1}^- \tag{5.4}$$

where the phase incrementing elements of the positive and negative sequence elements are complex conjugates of each other. Therefore, v_n can be expressed using the widely linear model given by

$$\begin{bmatrix} v_n \\ v_n^* \end{bmatrix} = \begin{bmatrix} v_{n-1}^+ & v_{n-1}^- \\ v_{n-1}^{-*} & v_{n-1}^{+*} \end{bmatrix} \begin{bmatrix} \varphi_n \\ \varphi_n^* \end{bmatrix} \tag{5.5}$$

where $\varphi_n = e^{i2\pi f\Delta T}$ represents the phase increment coefficient. A geometric interpretation of the output of the Clarke transform, v_n, for both a balanced and an unbalanced three-phase power system that is experiencing an 80% drop in the amplitude of $v_{a,n}$ and a 20 degree shift in the phases of $v_{b,n}$ and $v_{c,n}$ is shown in Figure 5.1; in addition, in Figure 5.2, the positive and negative sequence elements of the unbalanced three-phase power system are illustrated.

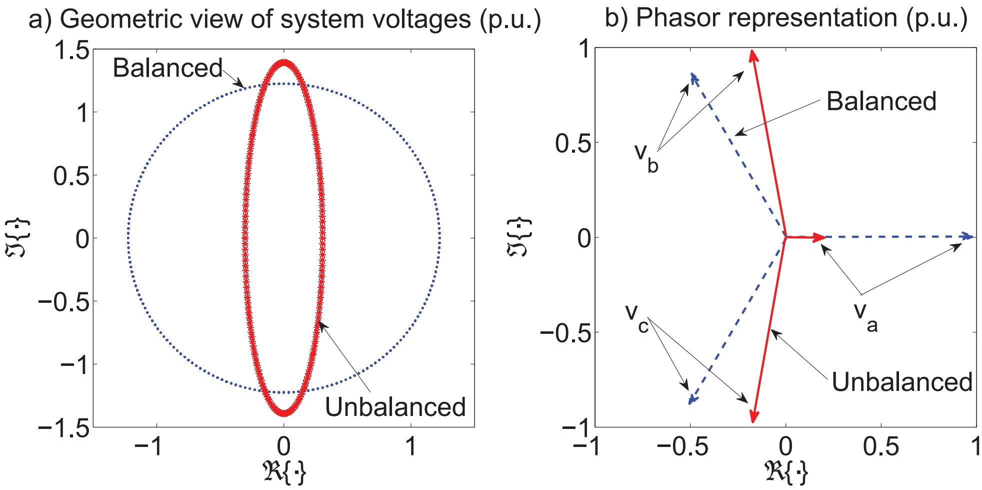

Figure 5.1: Geometric view of the output of the Clarke transform, v_n, for both a balanced and an unbalanced three-phase power system: a) geometric view of the output of the Clarke transform, b) phasor representation.

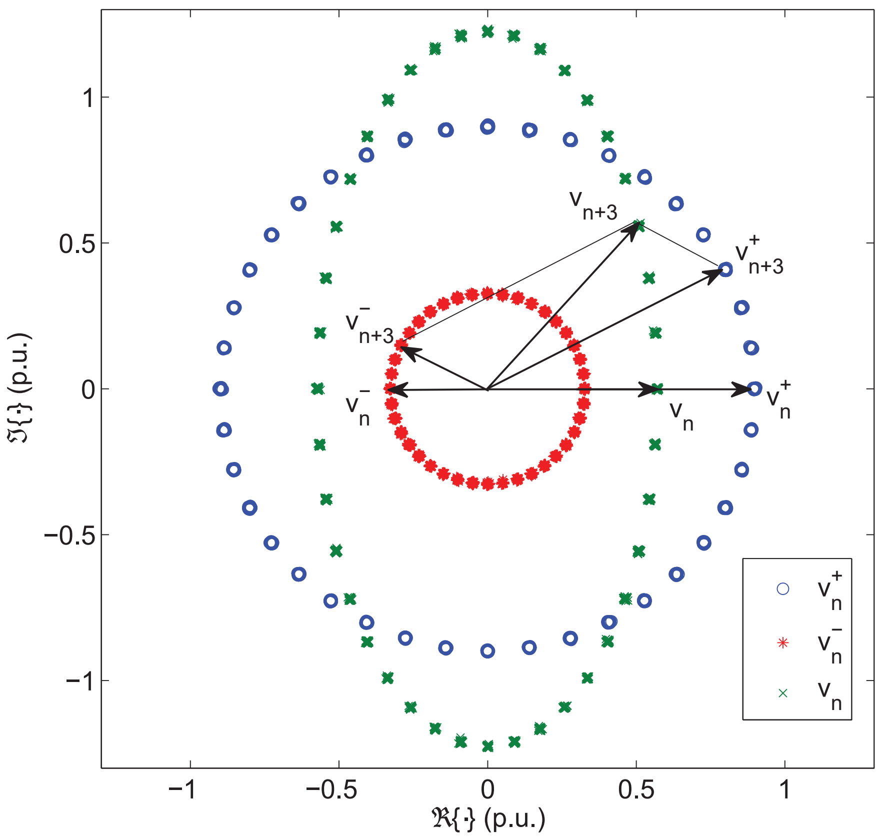

Figure 5.2: Geometrical view of the distribution of v_n in dark green "×", v_n^- in red "*", and v_n^+ in blue "o" for an unbalanced three-phase power system experiencing an 80% drop in the amplitude of $v_{a,n}$ and a 20 degree shift in the phases of $v_{b,n}$ and $v_{c,n}$. Note that $v_n = v_n^+ + v_n^-$.

Note that in the widely-linear model given in (5.5), the positive and negative sequence elements, v_n^+ and v_n^-, need to be known in order to estimate the phase incrementing element, φ_n. However, this is not the case as only observations of the output of the Clarke transform, v_n, are at hand. Although a dual Kalman filtering approach can be taken to simultaneously estimate φ_n and the system model parameters, v_n^+ and v_n^-, in order to simplify our approach and develop a more accurate and computationally efficient frequency estimator, by taking into account that the phase increment elements of v_n^+ and v_n^- are complex conjugates of each other, a widely-linear state space model for v_n is presented in Algorithm 6, where the fundamental frequency of the system can be estimated directly from the phase increment element, which is modeled as a state.

Algorithm 6. Complex Non-Linear Frequency Estimator (CNLFE)

State evolution equation:

$$\begin{bmatrix} \varphi_n \\ v_n^+ \\ v_n^- \\ \varphi_n^* \\ v_n^{+*} \\ v_n^{-*} \end{bmatrix} = \begin{bmatrix} \varphi_{n-1} \\ \varphi_{n-1} v_{n-1}^+ \\ \varphi_{n-1}^* v_{n-1}^- \\ \varphi_{n-1}^* \\ \varphi_{n-1}^* v_{n-1}^{+*} \\ \varphi_{n-1} v_{n-1}^{-*} \end{bmatrix} + \boldsymbol{\nu}_n^a$$

Observation equation:

$$\begin{bmatrix} v_n \\ v_n^* \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \varphi_n \\ v_n^+ \\ v_n^- \\ \varphi_n^* \\ v_n^{+*} \\ v_n^{-*} \end{bmatrix} + \boldsymbol{\omega}_n^a$$

Estimate of frequency:

$$\hat{f}_n = \frac{\Im\left(\ln\left(\varphi_n\right)\right)}{i 2\pi \Delta T}$$

5.5 Harmonic contamination

The ever increasing presence of loads with non-linear voltage-current characteristics, that is loads that draw a non-sinusoidal current from a sinusoidal voltage source such as compensators, inverters, direct current converters, and electric motor drives, results in distortion of the current and/or voltage signals from their desired sinusoidal shape [123]. In power engineering, these distorted voltage and current signals are mathematically dealt with through the framework of the Fourier theory that states: a periodic wave can be represented in terms of a summation of sinusoidal waves with frequencies at integer multiples of the fundamental frequency, referred to as harmonics. Since frequency estimators, such as the CLFE and CWLFE presented in Algorithm 4 and Algorithm 5, are designed for perfect sinusoidal wave forms, harmonic contamination is a major cause of error in these frequency estimators.

In a similar manner to that explained in Section 5.4 for the main component of the signal, the m^{th} harmonic component after passing through the Clarke transform can be expressed in the form of

$$v_{n,m} = \Lambda_{I,n,m}\cos(2\pi mf\Delta Tn) - \Lambda_{Q,n,m}\sin(2\pi mf\Delta Tn)$$

with $\{\Lambda_{I,n,m}, \Lambda_{Q,n,m}\} \in \mathbb{C}$ and given by

$$\begin{aligned}\Lambda_{I,n,m} =& \sqrt{\frac{2}{3}}V_{a,n,m}\cos(\phi_{a,n,m}) + (\frac{i\sqrt{3}-1}{\sqrt{6}})V_{b,n,m}\cos\big(\phi_{b,n,m} + \frac{2\pi}{3}\big)\\ &- (\frac{i\sqrt{3}+1}{\sqrt{6}})V_{c,n,m}\cos\big(\phi_{c,n,m} + \frac{4\pi}{3}\big)\\ \Lambda_{Q,n,m} =& \sqrt{\frac{2}{3}}V_{a,n,m}\sin(\phi_{a,n,m}) + (\frac{i\sqrt{3}-1}{\sqrt{6}})V_{b,n,m}\sin\big(\phi_{b,n,m} + \frac{2\pi}{3}\big)\\ &- (\frac{i\sqrt{3}+1}{\sqrt{6}})V_{c,n,m}\sin\big(\phi_{c,n,m} + \frac{4\pi}{3}\big).\end{aligned}$$

where $\{V_{c,n,m}, V_{c,n,m}, V_{c,n,m}\}$, denote the instantaneous amplitudes of the m^{th} harmonic voltages and $\{\phi_{a,n,m}, \phi_{a,n,m}, \phi_{a,n,m}\}$ are the instantaneous phase shifts of the m^{th} harmonic voltages.

In this section, two methods for dealing with harmonic contamination in three-phase power signals is presented. The first approach simply considered harmonics as noise components of the observed signals. The second approach takes advantage of the flexibility of the developed CNLFE, presented in Algorithm 6, to incorporate the main harmonic components as part of the signal.

5.5.1 Harmonics as noise

In most three-phase power distribution systems adequate provisions are put in place to reduce the spread of harmonics, due to their negative effect on the life span and efficiency of power grid components; in addition, the amount of harmonics that consumers are allowed to create is strictly regulated [106, 123]. Therefore, under conditions where the harmonic components of the signal is limited, it would be reasonable from a computational complexity point of view to simply account these harmonics as the noise component of the signal.

The covariance of the m^{th} harmonic component of three-phase power signal modeled as noise is given by

$$\begin{aligned} E\left[v_{n,m}v_{n,m}^{*}\right] =& E\left[\Lambda_{I,n,m}\cos(2\pi mf\Delta Tn)\Lambda_{I,n,m}^{*}\cos(2\pi mf\Delta Tn)\right] \\ &+ E\left[\Lambda_{Q,n,m}\sin(2\pi mf\Delta Tn)\Lambda_{Q,n,m}^{*}\sin(2\pi mf\Delta Tn)\right] \\ &- E\left[\Lambda_{I,n,m}\cos(2\pi mf\Delta Tn)\Lambda_{Q,n,m}^{*}\sin(2\pi mf\Delta Tn)\right] \\ &- E\left[\Lambda_{I,n,m}^{*}\cos(2\pi mf\Delta Tn)\Lambda_{Q,n,m}\sin(2\pi mf\Delta Tn)\right] \\ =& \frac{|\Lambda_{I,n,m}|^2}{2} + \frac{|\Lambda_{Q,n,m}|^2}{2} \end{aligned} \tag{5.6}$$

whereas its pseudo-covariance can be formulated as

$$\begin{aligned} E\left[v_{n,m}v_{n,m}\right] =& \frac{1}{2}E\left[\Lambda_{I,n,m}^2 + \Lambda_{I,n,m}^2\cos(4\pi mf\Delta Tn)\right] \\ &+ \frac{1}{2}E\left[\Lambda_{Q,n,m}^2 - \Lambda_{Q,n,m}^2\cos(4\pi mf\Delta Tn)\right] \\ &- 2E\left[\Lambda_{I,n,m}\cos(2\pi mf\Delta Tn)\Lambda_{Q,n,m}\sin(2\pi mf\Delta Tn)\right] \\ =& \frac{\Lambda_{I,n,m}^2}{2} + \frac{\Lambda_{Q,n,m}^2}{2}. \end{aligned} \tag{5.7}$$

From the expressions in (5.6) and (5.7), note that if the three-phase voltages of the m^{th} harmonic component are balanced; then

$$E\left[v_{n,m}v_{n,m}^{*}\right] = \frac{V_{a,n,m}^2 + V_{b,n,m}^2 + V_{c,n,m}^2}{2} \qquad \text{and} \qquad E\left[v_{n,m}v_{n,m}\right] = 0.$$

However, this is not the case when the three-phase voltages of the m^{th} harmonic are not balanced, leading to an improper observational noise[1]. This furtherer validates our approach in using complex-valued widely-linear modeling as the information in the pseudo-covariance cannot be exploited through standard linear complex-valued modeling of the three-phase power system.

5.5.2 Harmonics as part of the signal

Although, from a computational complexity point of view, it is convenient to model harmonic components of the signal as noise, this approach compromises the performance of frequency estimators as the energy of the section of the signal considered as observational noise rises rapidly with increase in harmonic contamination. In addition, it is important in harmonic mitigation applications to be able to account for the presence of the major harmonic components. Therefore, it is prudent to model the major harmonic components as part of the signal.

[1] It is most likely that the three-phase voltages of the major harmonic components, such as the third and fifth-order harmonics, be unbalanced, even if the system is operating under balanced conditions [123].

In contrast to the state space parameters in Algorithm 5, the state space parameters in Algorithm 6 have an intuitive physical interpretation, which allows for the framework to be extended to include the main harmonic components in the three-phase power system if necessary. For instance, a state space model taking into account a harmonic at m times the main frequency of the system is given in Algorithm 7, where $v^+_{n,m}$ and $v^-_{n,m}$ denote the positive and negative sequences of the m^{th} harmonic component.

Algorithm 7. CNLFE With Harmonic Contamination (CNLFE-WHC)

State evolution equation:

$$\begin{bmatrix} \varphi_n \\ v^+_n \\ v^-_n \\ v^+_{n,m} \\ v^-_{n,m} \\ \varphi^*_n \\ v^{+*}_n \\ v^{-*}_n \\ v^{+*}_{n,m} \\ v^{-*}_{n,m} \end{bmatrix} = \begin{bmatrix} \varphi_{n-1} \\ \varphi_{n-1} v^+_{n-1} \\ \varphi^*_{n-1} v^-_{n-1} \\ \varphi^m_{n-1} v^+_{n-1,m} \\ \varphi^{*m}_{n-1} v^-_{n-1,m} \\ \varphi^*_{n-1} \\ \varphi^*_{n-1} v^{+*}_{n-1} \\ \varphi_{n-1} v^{-*}_{n-1} \\ \varphi^{*m}_{n-1} v^{+*}_{n-1,m} \\ \varphi^m_{n-1} v^{-*}_{n-1,m} \end{bmatrix} + \boldsymbol{\nu}^a_n$$

Observation equation:

$$\begin{bmatrix} v_n \\ v^*_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \varphi_n \\ v^+_n \\ v^-_n \\ v^+_{n,m} \\ v^-_{n,m} \\ \varphi^*_n \\ v^{+*}_n \\ v^{-*}_n \\ v^{+*}_{n,m} \\ v^{-*}_{n,m} \end{bmatrix} + \boldsymbol{\omega}^a_n$$

Estimate of frequency:

$$\hat{f}_n = \frac{\Im\left(\ln\left(\varphi_k\right)\right)}{i2\pi\Delta T}$$

Note that the CNLFE-WHC frequency estimator presented in Algorithm 7 requires more processing power as compared to that of the CNLFE presented in Algorithm 6; however, it accounts for the presence of the m^{th} harmonic allowing the measure, $\kappa_m = \left|v^+_{n,m}\right|^2 + \left|v^+_{n,m}\right|^2$, to be used to detect its presence. In addition, considering the m^{th} harmonic as part of the signal provides for a lower variance observational noise and better estimates of the main frequency of the system.

5.6 Performance analysis

In terms of computational complexity, the newly developed CNLFE presented in Algorithm 6 and its counterpart, the CWLFE presented in Algorithm 5, are implemented using the AECKF, have state vectors of similar lengths, and have identical observational vectors; however, the CWLFE in Algorithm 5 estimates the system frequency through a complicated function of the state vector elements, resulting in higher computational complexity. In addition, the CWLFE can only model harmonic components of the signal as observational noise, whereas it was shown in Section 5.5.2 that in the framework of the CNLFE, the flexibility is available to model the harmonic components of the signal as either observational noise or as part of the signal, dependent on the application and processing power available. In this section, the performance of the CNLFE developed here and presented in Algorithm 6 is quantified and compared to that of its counterpart, the CWLFE presented in Algorithm 5, as a bench mark.

The AECKF filter provides an unbiased estimate of the augmented state vector along side an estimate of the augmented covariance matrix of the state vector estimation error. In this setting, the estimate of the phase incrementing element of the CNLFE can be modeled as

$$\hat{\varphi_n} = \varphi_n + u_n$$

where $\hat{\varphi_n}$ is the estimate of the phase incrementing element at time instant n and u_n is a zero-mean complex-valued Gaussian random variable representing the estimation error, the second-order statistics of which are also estimated by the AECKF. The estimate of the system frequency obtained by the CNLFE can now be expressed as

$$\hat{f}_n = \frac{\Im\left(\ln\left(\hat{\varphi}_k\right)\right)}{i2\pi\Delta T}.$$

Making the assumptions that the sampling interval, ΔT, is small enough to ensure $\Re(\varphi_n) = \cos(2\pi f\Delta T) > \Re(u_n)$, results in $\ln(\cdot)$ being situated in its analytical region. Furthermore, assuming that $|\varphi_n| = 1 \gg |u_n|$, allows $\ln(\cdot)$ to be approximated by its first-order Taylor expansion around φ_n which yields

$$\begin{aligned}\hat{f}_n =& \frac{\Im\left(\ln(\varphi_n + u_n)\right)}{i2\pi\Delta T} = \frac{1}{i2\pi\Delta T}\Im\left(\ln(\varphi_n) + \frac{u_n}{\varphi_n}\right)\\ =& \frac{\Im\left(\ln(\varphi_n)\right)}{i2\pi\Delta T} + \frac{1}{i2\pi\Delta T}\Im\left(\frac{u_n}{\varphi_n}\right) = f + \frac{1}{i2\pi\Delta T}\Im\left(\frac{u_n}{\varphi_n}\right).\end{aligned} \tag{5.8}$$

Taking into account that $E[u_n] = 0$, the statistical expectation of $\hat{f}_n$, as expressed in (5.8), is given by

$$E[\hat{f}_n] = f + \frac{1}{i2\pi\Delta T}E\left[\Im\left(\frac{u_n}{\varphi_n}\right)\right] = f + \frac{1}{i2\pi\Delta T}\Im\left(\frac{E[u_n]}{u_n}\right) = f \tag{5.9}$$

indicating that the developed CNLFE produces unbiased estimates of the system frequency. In addition, the MSE of the estimates of the system frequency can now be formulated as

$$E\left[\left(\hat{f}_n - f\right)^2\right] = E\left[\frac{-1}{4(\pi\Delta T)^2}\Im^2\left(\frac{u_n}{\varphi_n}\right)\right] = \frac{-1}{8\left(\pi\Delta T\right)^2}\left(\Re\left(\frac{R_{u_n u_n}}{\varphi_n^2}\right) - C_{u_n}\right)$$

where replacing $\varphi_n^2 = \cos(4\pi f\Delta T) + \imath\sin(4\pi f\Delta T)$ yields

$$\begin{aligned} E\left[\left(\hat{f}_n - f\right)^2\right] =& \frac{-1}{8\left(\pi\Delta T\right)^2}\left(\Re\left(\frac{R_{u_n u_n}}{\varphi_n^2}\right) - C_{u_n}\right) \\ =& \frac{-1}{8\left(\pi\Delta T\right)^2}\Big(\cos(4\pi f\Delta T)\Re(R_{u_n u_n}) + \sin(4\pi f\Delta T)\Im(R_{u_n u_n}) - C_{u_n}\Big). \end{aligned} \tag{5.10}$$

Now, taking into account that the system frequency in three-phase power applications is usually close to 50 Hz or 60 Hz whereas the sampling frequency is on the order of ten to twenty times larger, it is reasonable to assume $\cos(4\pi f\Delta T) \approx 1$ and $\sin(4\pi f\Delta T) \approx 0$; therefore, from the expression in (5.10) we have

$$E\left[\left(\hat{f}_n - f\right)^2\right] = \frac{1}{8\left(\pi\Delta T\right)^2}\Big(C_{u_n} - \Re(R_{u_n u_n})\Big) = -\frac{E\left[\Im^2(u_n)\right]}{4\left(\pi\Delta T\right)^2} = \frac{C_{\Im(u_n)}}{4\left(\pi\Delta T\right)^2}. \tag{5.11}$$

Note that from the expressions in (5.9) and (5.11) it becomes apparent that the developed CNLFE is unbiased and its steady-state MSE is not directly dependent on the operating conditions of the three-phase system.

Recall that the estimate of the system frequency at time instant n obtained by the CWLFE is given by

$$\hat{f}_n = \frac{\arcsin\left(-i\Im\left(h_n + a_n\right)\right)}{2\pi\Delta T} \tag{5.12}$$

with

$$a_n = -\Im\left(h_n\right) + \sqrt{\Im^2\left(h_n\right) + |g_n|^2} \tag{5.13}$$

while replacing the expression in (5.13) into the expression in (5.12) yields

$$\hat{f}_n = \frac{1}{2\pi\Delta T}\arcsin\left(-i\sqrt{\Im^2(h_n) + |g_n|^2}\right)$$

where it is important to note that $-i\sqrt{\Im^2(h_n) + |g_n|^2} \in \mathbb{R}^+$. The estimates of h_n and g_n obtained by the AECKF can be modeled as

$$\begin{aligned} \hat{h}_n =& h_n + u_{h_n} \\ \hat{g}_n =& g_n + u_{g_n} \end{aligned}$$

where $\hat{h}_n$ and $\hat{g}_n$ are estimates of h_n and g_n, while u_{h_n} and u_{g_n} are zero-mean complex-valued Gaussian random variables representing the estimation error of h_n and g_n respectively. Therefore, the estimate of the system frequency obtained by the CWLFE can be expressed as

$$\begin{aligned}\hat{f}_n =& \frac{1}{2\pi\Delta T}\arcsin\left(-i\sqrt{\Im^2(h_n+u_{h_n})+|g_n+u_{g_n}|^2}\right)\\ =& \frac{1}{2\pi\Delta T}\arcsin\left(-i\sqrt{\Im^2(h_n)+\Im^2(u_{h_n})+2\Im(h_n)\Im(u_{h_n})+|g_n|^2+|u_{g_n}|^2+2\Re(g_n u_{g_n})}\right)\end{aligned}$$

where assuming that $|h_n| \gg |u_{h_n}|$ and $|g_n| \gg |u_{g_n}|$ and approximating the square root function using its first-order Taylor expansion around $\Im^2(h_n)+|g_n|^2$ yields

$$\hat{f}_n = \frac{1}{2\pi\Delta T}\arcsin\left(-i\sqrt{\Im^2(h_n)+|g_n|^2}+\frac{\Im^2(u_{h_n})+2\Im(h_n)\Im(u_{h_n})+|u_{g_n}|^2+2\Re(g_n u_{g_n})}{2i\sqrt{\Im^2(h_n)+|g_n|^2}}\right).$$

Now, approximating $\arcsin(\cdot)$ around $-i\sqrt{\Im^2(h_n)+|g_n|^2}$ with its first-order Taylor series expansion allows the estimate of the system frequency to be expressed as

$$\begin{aligned}\hat{f}_n =& \frac{1}{2\pi\Delta T}\arcsin\left(-i\sqrt{\Im^2(h_n)+|g_n|^2}\right)\\ &+\frac{\Im^2(u_{h_n})+2\Im(h_n)\Im(u_{h_n})+|u_{g_n}|^2+2\Re(g_n u_{g_n})}{i4\pi\Delta T\sqrt{1-\Im^2(h_n)-|g_n|^2}\sqrt{\Im^2(h_n)+|g_n|^2}}\\ =& f+\frac{\Im^2(u_{h_n})+2\Im(h_n)\Im(u_{h_n})+|u_{g_n}|^2+2\Re(g_n u_{g_n})}{i4\pi\Delta T\sqrt{1-\Im^2(h_n)-|g_n|^2}\sqrt{\Im^2(h_n)+|g_n|^2}}.\end{aligned} \tag{5.14}$$

The statistical expectation of $\hat{f}_n$, as expressed in (5.14), is now given by

$$E[\hat{f}_n] = f+\frac{E\left[\Im^2(u_{h_n})\right]+2E\left[\Im(h_n)\Im(u_{h_n})\right]+E\left[|u_{g_n}|^2\right]+2E\left[\Re(g_n u_{g_n})\right]}{4\pi\Delta T\sqrt{1-\Im^2(h_n)-|g_n|^2}\left(i\sqrt{\Im^2(h_n)+|g_n|^2}\right)}$$

where considering that that u_{h_n} and u_{g_n} are zero-mean yields[2]

$$E[\hat{f}_n] = f+\frac{C_{u_{g_n}}-C_{\Im(u_{h_n})}}{4\pi\Delta T\sqrt{1-\Im^2(h_n)-|g_n|^2}\left(i\sqrt{\Im^2(h_n)+|g_n|^2}\right)}. \tag{5.15}$$

Therefore, the CWLFE cannot guarantee unbiased estimates of the system frequency regardless of the operating conditions of the three-phase system.

[2]Once again, note that $i\sqrt{\Im^2(h_n)+|g_n|^2}\in\mathbb{R}$; therefore, $E[\hat{f}_n]$ as expressed in (5.15) lies in $\mathbb{R}$.

The MSE of the estimate of the system frequency, as expressed in (5.14), can be formulated as

$$E\left[\left(f-\hat{f}_n\right)^2\right]=\frac{E\left[\left(\Im^2(u_{h_n})+2\Im(h_n)\Im(u_{h_n})+|u_{g_n}|^2+2\Re(g_n u_{g_n})\right)^2\right]}{16\left(\pi\Delta T\right)^2\left(1-\Im^2(h_n)-|g_n|^2\right)\left|\Im^2(h_n)+|g_n|^2\right|}. \tag{5.16}$$

Now, assuming that the sampling frequency is much higher than the system frequency results in $\left(1-\Im^2(h_n)-|g_n|^2\right)\approx 1$; furthermore, considering the terms $\Im^2(u_{h_n})$ and $|u_{g_n}|^2$ to be small compared to $2\Im(h_n)\Im(u_{h_n})$ and $2\Re(g_n u_{g_n})$ allows the expression in (5.16) to be simplified into

$$E\left[\left(f-\hat{f}_n\right)^2\right]=\frac{E\left[\Im^2(h_n)\Im^2(u_{h_n})\right]+E\left[\Re^2(g_n u_{g_n})\right]}{4\left(\pi\Delta T\right)^2\left|\Im^2(h_n)+|g_n|^2\right|}. \tag{5.17}$$

Regarding the expressions in (5.16) and (5.17), it is important to make the following three remarks:

1. The expressions in (5.16) and (5.17) are dependent on h_n and g_n and hence on the operating conditions of the three-phase system.

2. For a constant h_n, the expressions in (5.16) and (5.17) are minimized when $g_n=0$, which gives

$$E\left[\left(f-\hat{f}_n\right)^2\right]\geq\frac{E\left[\left(\Im^2(u_{h_n})+2\Im(h_n)\Im(u_{h_n})+|u_{g_n}|^2\right)^2\right]}{16\left(\pi\Delta T\right)^2\left(1-\Im^2(h_n)\right)\left|\Im^2(h_n)\right|}\approx\frac{E\left[\Im^2(u_{h_n})\right]}{4\left(\pi\Delta T\right)^2}. \tag{5.18}$$

3. From comparing the expressions in (5.11) and (5.18), it becomes apparent that the CNLFE achieves a smaller steady-state MSE than that of CWLFE, that is assuming both algorithms achieve the same accuracy on their state vector estimates or $E\left[\Im^2(u_n)\right]\approx E\left[\Im^2(u_{h_n})\right]$.

5.7 Simulations

In this section, the performance of the newly developed CNLFE is validated and compared to that of the CLFE and the CWLFE in different experiments involving practical power grid scenarios. In all experiments the sampling frequency was $f_s=1$ kHz and the voltage measurements were considered to be corrupted by white Gaussian noise with SNR of 40 dB.

5.7.1 Synthetically generated data

In the first experiment, the three-phase system was considered to be initially balanced, then suffered a fault resulting in a voltage sag characterized by an 80% drop in the amplitude of $v_{a,n}$ and 20 degree shifts in the phases of $v_{b,n}$ and $v_{c,n}$, the voltage phasor plot of which is shown in Figure 5.1. Furthermore, the frequency of the system experienced a step jump of 2 Hz. The fault lasted for a short duration and the system returned to balanced operating conditions and its nominal frequency once more. The estimates of the system frequency obtained through employing the CLFE, CWLFE, and CNLFE are shown in Figure 5.3. Observe that the CLFE could only accurately estimate the system frequency during balanced operating conditions whereas the CWLFE and CNLFE were able to track the system frequency during both balanced and unbalanced operating conditions. Moreover, a closer examination of the transient behavior of the CWLFE and CNLFE is made in Figure 5.4. Notice that both when the three-phase system was starting to experience the voltage sag and was recovering from the voltage sag, the CNLFE outperformed the CWLFE in terms of convergence rate and had a better dynamic behavior, that is the CNLFE had less overshoot and undershoot.

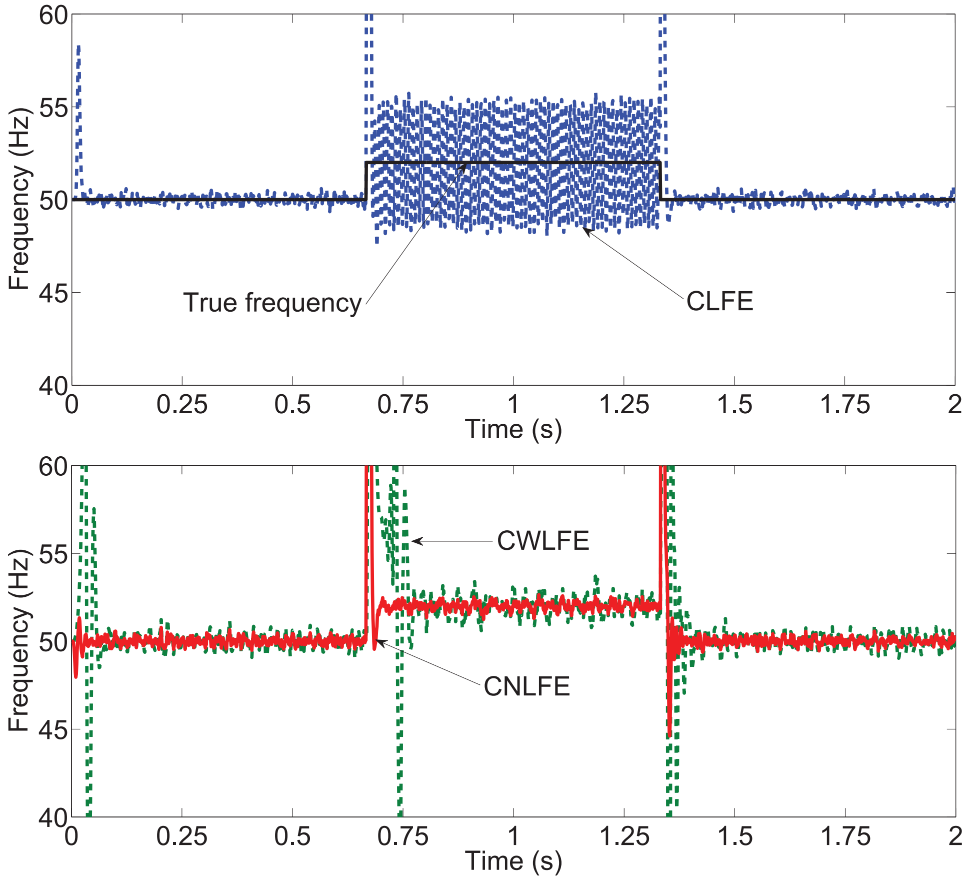

Figure 5.3: Frequency estimation for a three-phase power system experiencing a voltage sag and a 2 Hz jump in frequency from 0.667 to 1.334 seconds. The performance of the CLFE is shown in the top graph while the performance of the CNLFE and CWLFE are compared in the bottom graph.

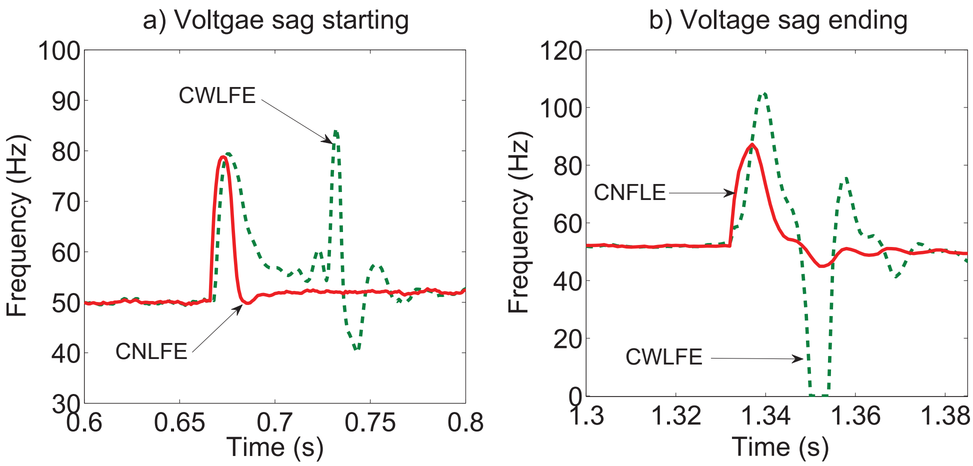

Figure 5.4: Transient behavior of the CWLFE and CNLFE when the three-phase power system experiences a voltage sag and a 2 Hz step jump in frequency: a) voltage sag starting, b) voltage sag ending.

In order to further validate the performance of the CNLFE, in the next experiment the unbalanced three-phase system that is characterized in Figure 5.1 was considered to experience a rising (*cf.* falling) frequency due to mismatch between power generation and consumption 0.5 seconds after the simulation started. The estimates of the system frequency obtained through employing the CNLFE and CWLFE are shown in Figure 5.5. Observe that both when the system frequency was constant and the system frequency was changing, the CWLFE and CNLFE accurately tracked the system frequency, with the CNLFE outperforming the CWLFE in terms of steady-state variance.

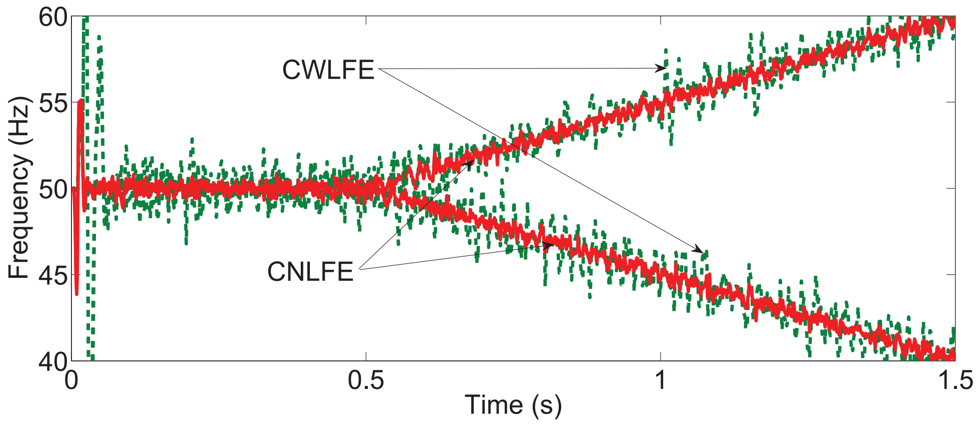

Figure 5.5: Frequency estimation for an unbalanced three-phase power system experiencing a changing frequency at the rate of 10 Hz/s employing the CWLFE and CNLFE.

The steady-state MSE performance advantage of the newly developed CNLFE for various SNR values is shown in Figure 7.7, where the steady-state MSE of the CLFE, CWLFE, and CNLFE for both a balanced and the unbalanced three-phase system shown in Figure 5.1 are compared. Notice that the developed CNLFE achieves a better MSE performance than the CWLFE and CLFE, verifying the analysis in Section 5.6. Furthermore, unbalanced operating conditions did not have an effect on the steady-state MSE performance of the CNLFE, a desirable characteristic for frequency estimators of three-phase power systems and in agreement with the analysis in Section 5.6.

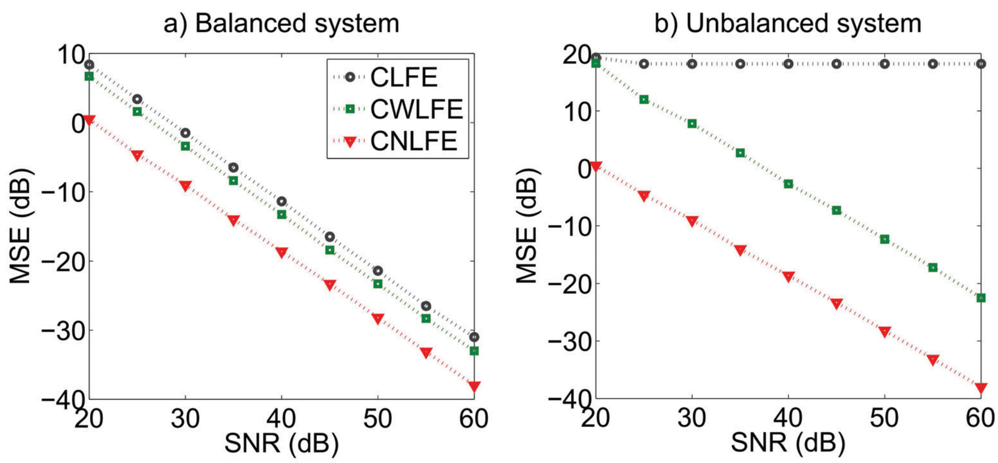

Figure 5.6: Steady-state MSE performance analysis for SNR ranging from 20 dB to 60 dB: a) balanced three-phase system, b) unbalanced three-phase system.

The steady-state bias performance advantage of the newly developed CNLFE for various SNR values is now investigated. In Figure 5.7, the steady-state bias performance of the CNLFE is compared to that of the CWLFE and the CLFE for a balanced and the unbalanced three-phase system characterized in Figure 5.1. Observe that the developed CNLFE achieves a lower bias than the CWLFE and CLFE, verifying the analysis in Section 5.6. It is important to note that for the unbalanced three-phase system the CLFE produced estimates with around 1.7 Hz of bias regardless of the SNR of the voltage measurements.

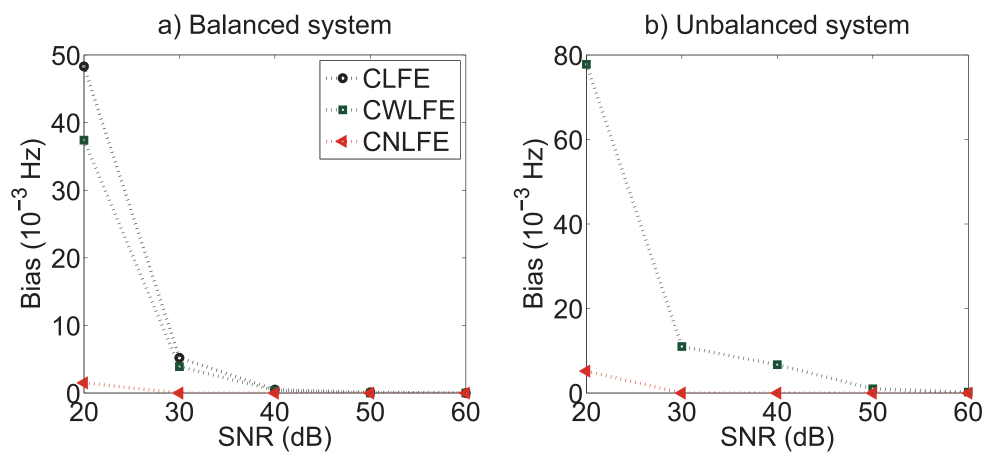

Figure 5.7: Steady-state bias performance analysis for SNR ranging from 20 dB to 60 dB: a) balanced three-phase system, b) unbalanced three-phase system.

In the next experiment, the effect of harmonics on the performance of the CNLFE is considered. This is performed through the addition of a balanced 10% third harmonic component to the unbalanced three-phase system used in the previous experiment at $t = 0.05$ s. The system voltages and estimates of the system frequency are shown in Figure 5.8. Notice that the proposed CNLFE achieved a smaller steady-state oscillatory error than the CWLFE; furthermore, when the proposed CNLFE was used in its modified format, CNLFE-WHC, the frequency of the system was estimated accurately regardless of the presence of the third harmonic component.

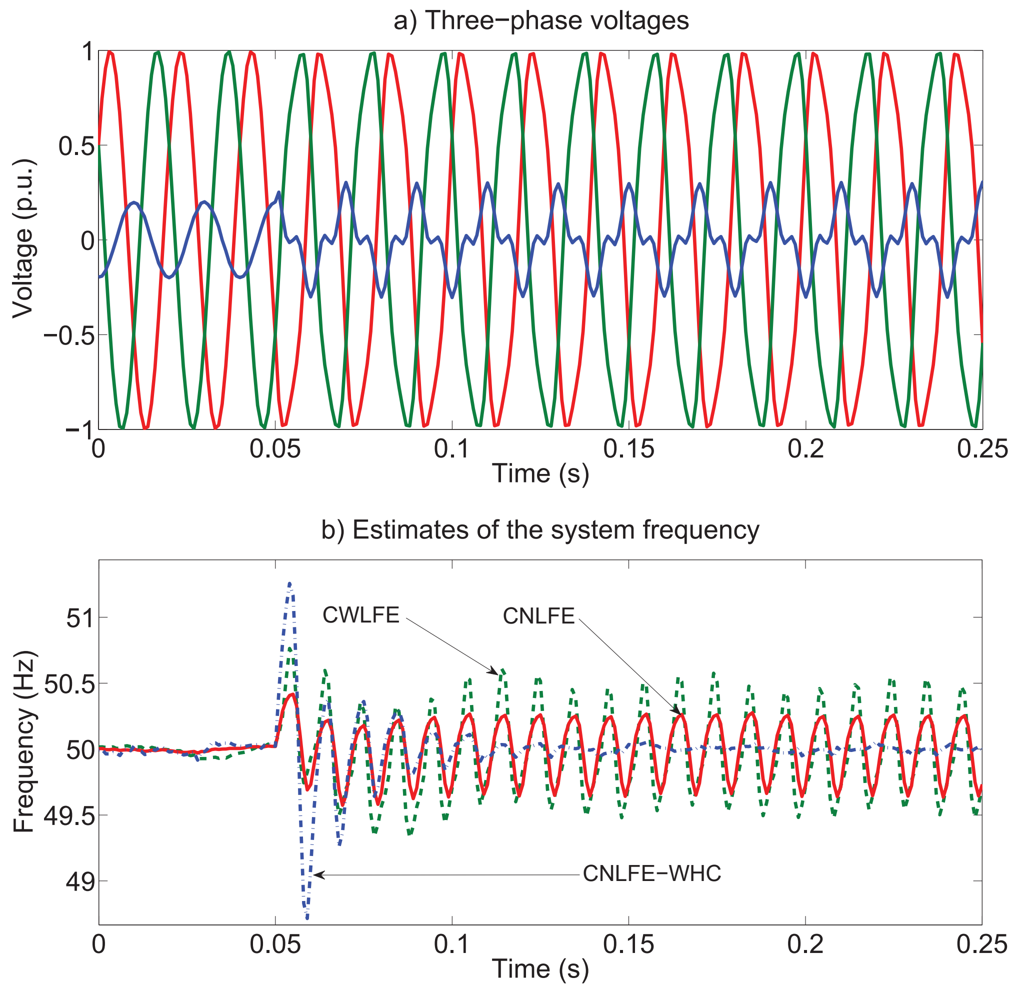

Figure 5.8: Frequency estimation for an unbalanced three-phase power system contaminated with a 10% third order harmonic: a) three-phase voltages, b) estimates of the system frequency.

5.7.2 Real-world data

The performance of the developed CNLFE is now evaluated in a more practical setting using real-world data recorded at a 110/20/10 k.V. transformer station. The REL 531 numerical line distance protection terminal, produced by ABB Ltd, was installed in the station and was used to record the "phase-ground" voltages.

In the first simulation using real-world data, the three-phase system appeared to be operating in a balanced fashion for the first 5 seconds, then experienced a severe voltage sag at approximately 5.05 seconds after recoding started which lasted for around 80 milliseconds. The recorded data and the estimates of the system frequency are shown in Figure 5.9. Observe the CLFE and the CWLFE lost track of the system frequency during the voltage sag whereas the CNLFE was able to track the system frequency during the voltage sag and showed outstanding performance.

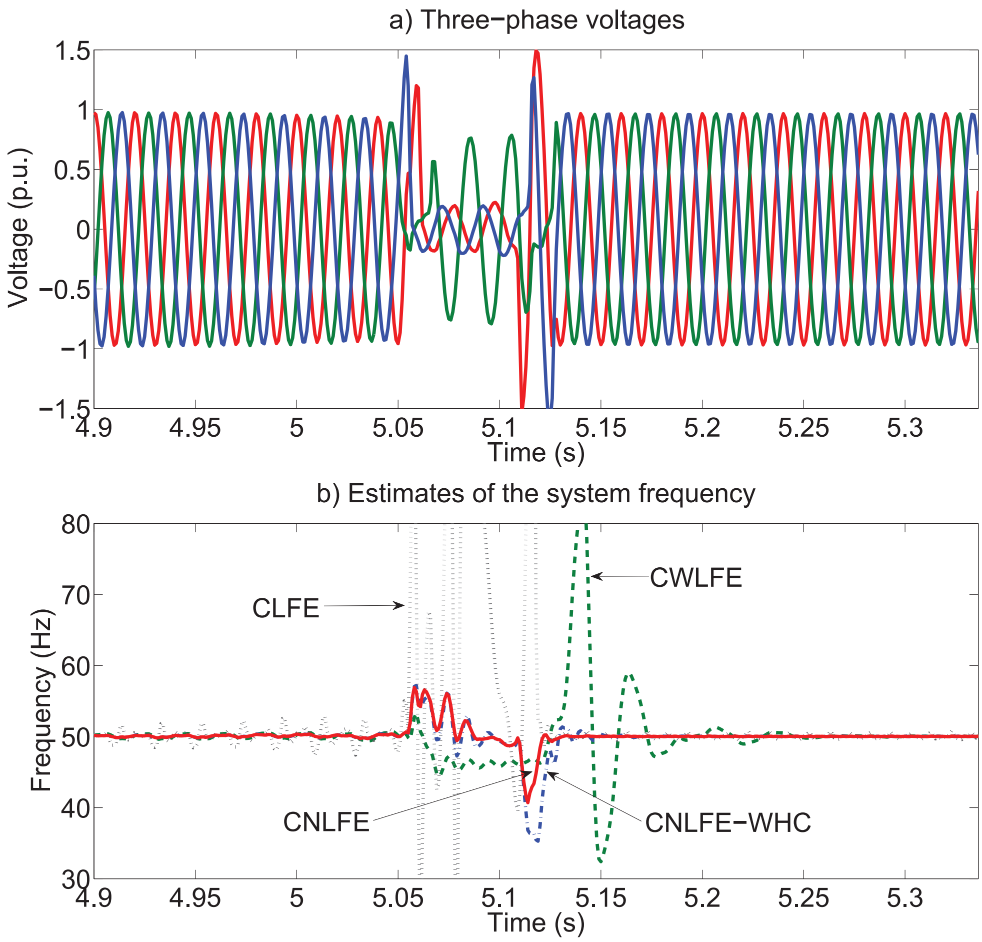

Figure 5.9: Frequency estimation during a sever real-world voltage sag: a) three-phase voltages, b) estimates of the system frequency.

In the second simulation using real-world data, the three-phase system suffered a voltage sag approximately 2.5 seconds after recording started. The recorded data and the estimates of the system frequency are shown in Figure 5.10. Notice that once again the newly developed CNLFE achieved a better steady-state variance as compared to the CWLFE and CLFE.

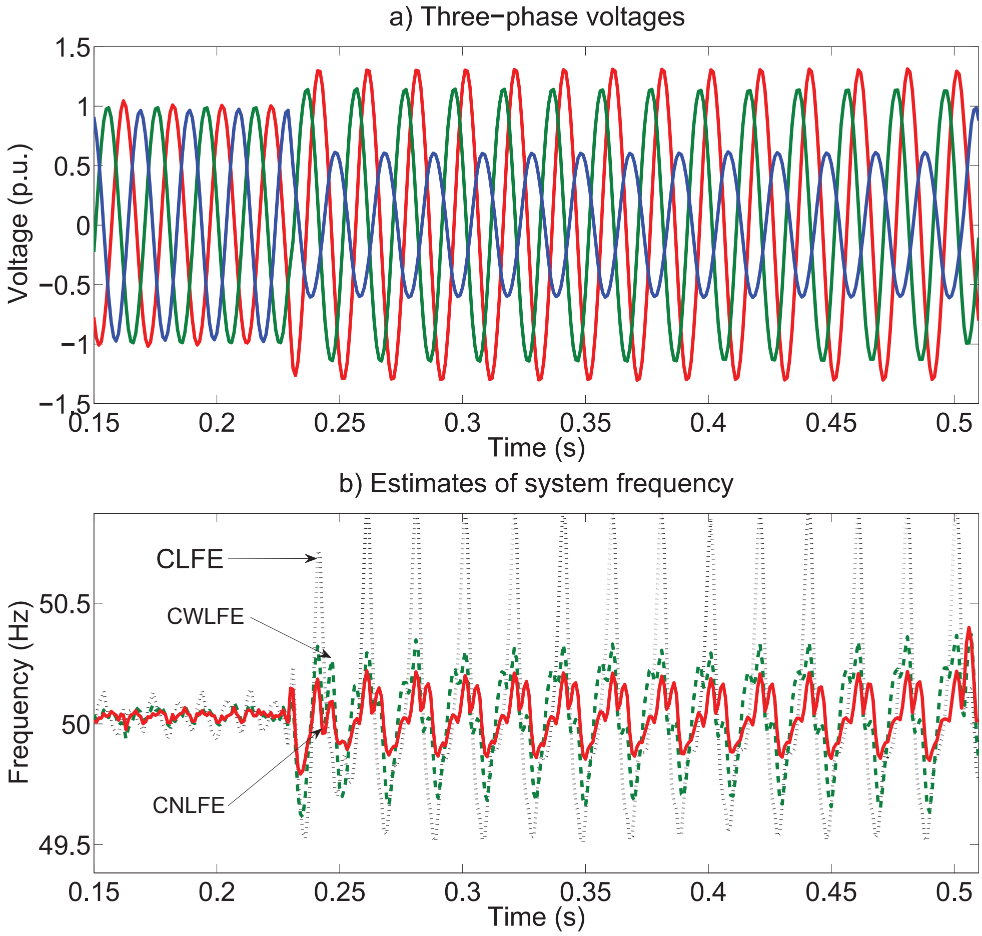

Figure 5.10: Frequency estimation during a voltage sag using real-world data recording: a) three-phase voltages, b) estimates of the system frequency.

5.8 Conclusion

Frequency estimation in three-phase power systems has been considered and a unified approach for estimating the fundamental frequency of both balanced and unbalanced three-phase systems based on the positive and negative sequence elements in the output of the Clarke transform has been presented. The output of the Clarke transform consists of only a positive sequence when the three-phase system is balanced; however, when the three-phase system is unbalanced, a negative sequence is also present in the signal, which compromises the performance of standard strictly linear estimators. The newly developed CNLFE accounts for the presence of the negative sequence element in the output of the Clarke transform and operates optimally under both balanced and unbalanced operating conditions. Furthermore, it has also been demonstrated that the framework can be easily extended to account for the presence of harmonics. The performance of the developed algorithm has been quantified and extensively tested using synthetic and real-world data where it was shown to outperform the strictly linear CLFE and the recently introduced widely-linear CWLFE, especially during critical moments when the three-phase system was experiencing a fault and/or harmonics.

Chapter 6

A Quaternion Joint Frequency and Phasor Estimator for Three-Phase Power Systems

6.1 Overview

In this chapter, a novel frequency estimator for three-phase power systems based on quaternion-valued state space modeling of three-phase power system signals is developed. Modeling the voltage measurements from all the phases of a three-phase power system as a quaternion-valued signal allows for the full characterization of the three-phase power systems. In addition, the components of the state space model are designed in a manner that permits the extraction of the voltage phasor information of each phase. To this end, an approach for estimating the voltage phasor information of three-phase power systems is also developed. For rigor, the performance of the developed frequency estimator is quantified and a framework is introduced for dealing with harmonic contamination.

6.2 Introduction

The need for rigorous frequency estimation techniques in power distribution systems was discussed in Section 5.2; however, this need becomes even more pronounced when considering current trends in smart grid technology that in addition to having to cater for an ever increasing unpredictable power consumption, incorporate distributed power generation based on renewable energy sources, such as solar and wind [57, 121, 122]. In this setting, the wide-area grid is divided into a number of self contained sections called micro-grids, with some micro-grids becoming independent in power generation and disconnecting from the wide-area grid for prolonged lengths of time, referred to as islanding. Perfect synchrony in frequency and voltage phasors is required to connect micro-grids and manage islanding; consequently, many smart grid control and management techniques are dependent on accurate estimation of frequency and voltage phasors under both balanced and unbalanced operating conditions.

For more than 50 years, the Clarke transform has formed the backbone of three-phase power system analysis, providing a complex-valued representation of three-phase power systems that allows for the use of the well defined complex-valued mathematical framework for analyzing the behavior of the three-phase power grid [53]. However, as was explained in Section 5.3, the Clarke transform can only truly represent all the information in a three-phase power system that is operating in a balanced fashion, as in essence the Clarke transform is a two-dimensional representation of a signal that is by nature three-dimensional. To this end, quaternions are employed in order to model the voltage measurements of thee-phase power systems directly in the three-dimensional space where they naturally reside, mitigating the need to use the Clarke transform.

In this work, a novel frequency estimator for three-phase power systems is derived based on quaternion-valued state space adaptive filtering. Furthermore, an insight to the physical interpretation of the elements of the state space vector is provided and exploited to estimate the voltage phasors of each phase. The resulting estimator can fully characterize three-phase systems, operates optimally under both balanced and unbalanced conditions, and eliminates the need to use the Clarke transform. Finally the performance of the developed algorithm is quantified and the developed algorithm is validated through simulation using both synthetic data and real-world data recordings where it is shown that it can outperform its complex-valued counterparts.

6.3 The quaternion frequency estimator

The three-phase voltages are combined together to generate a quaternion-valued signal with a vanishing real component given by

$$q_n = iv_{a,n} + jv_{b,n} + kv_{c,n} \tag{6.1}$$

where all the elements of q_n have the same frequency. Therefore, analytical geometry dictates that q_n traces an ellipse in a two-dimensional subspace (one plane) of the three-dimensional imaginary subspace of $\mathbb{H}$. This is shown in Figure 6.1, where the system voltages of a balanced and an unbalanced three-phase system are presented alongside a plot of their voltage phasors.

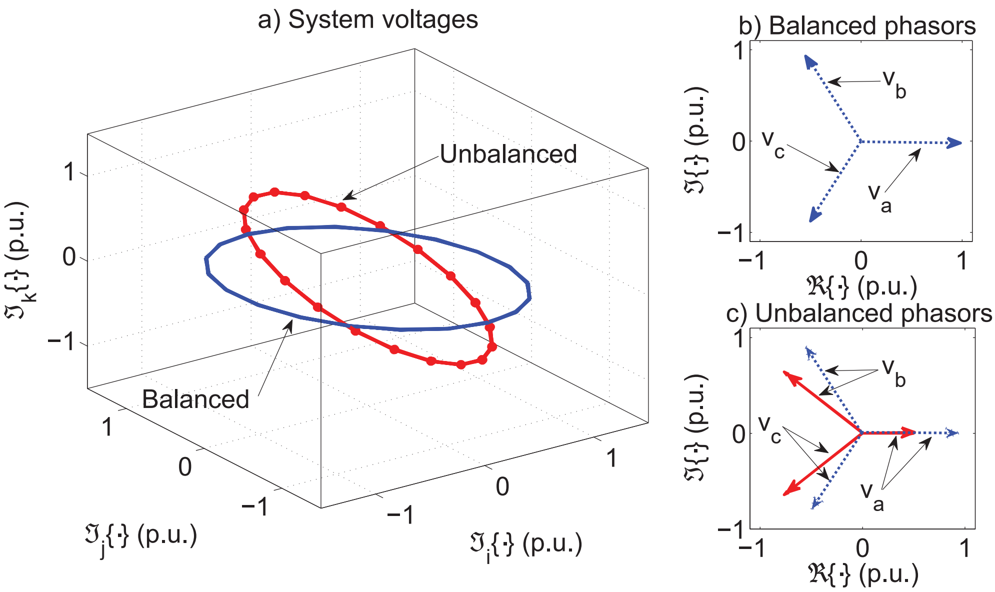

Figure 6.1: Geometric view of the system voltages, q_n, and the corresponding phasor diagrams of a balanced and an unbalanced three-phase system: a) system voltages, b) phasor representation of a balanced three-phase system, c) phasor representation of an unbalanced three-phase system. Solid red lines represent an unbalanced system, while dashed blue lines represent a balanced system.

In order to simplify our analysis and without loss of generality a new set of imaginary units, $\{\zeta, \zeta', \zeta''\}$, are defined such that

$$\zeta\zeta' = \zeta'', \;\; \zeta'\zeta'' = \zeta, \;\; \zeta''\zeta = \zeta'. \tag{6.2}$$

The ζ' and ζ'' imaginary units are designed to reside in the same plane as q_n, which results in ζ being normal to this plane. An arbitrary ellipse in the ζ'-ζ'' plane can then be expressed as

$$q_n = \zeta' A_n \text{sin}(2\pi f \Delta T n + \phi_{\zeta',n}) + \zeta'' B_n \text{sin}(2\pi f \Delta T n + \phi_{\zeta'',n}) \tag{6.3}$$

where $\{A_n, B_n\} \in \mathbb{R}$, are instantaneous amplitudes and $\{\phi_{\zeta',n}, \phi_{\zeta'',n}\} \in [0, 2\pi)$ are instantaneous phases. The expression in (6.3) can be rearranged using the expressions in (6.2) to give

$$q_n = \left(A_n \text{sin}\left(2\pi f \Delta T n + \phi_{\zeta',n}\right) + \zeta B_n \text{sin}\left(2\pi f \Delta T n + \phi_{\zeta'',n}\right)\right) \zeta'. \tag{6.4}$$

Given that $\zeta^2 = -1$, upon replacing the $\sin(\cdot)$ and $\cos(\cdot)$ functions with their polar representations, as given in (2.11), the expression in (6.4) yields

$$q_n = \left(e^{\zeta(2\pi f \Delta T n + \phi_{\zeta',n})} - e^{-\zeta(2\pi f \Delta T n + \phi_{\zeta',n})}\right)\left(\frac{A_n}{2\zeta}\right)\zeta' + \left(e^{\zeta(2\pi f \Delta T n + \phi_{\zeta'',n})} - e^{-\zeta(2\pi f \Delta T n + \phi_{\zeta'',n})}\right)\left(\frac{B_n}{2}\right)\zeta'. \tag{6.5}$$

Factoring out the terms $e^{\zeta(2\pi f \Delta T n)}$ and $e^{-\zeta(2\pi f \Delta T n)}$, the expression in (6.5) can be rearranged to give

$$q_n = \underbrace{e^{\zeta(2\pi f \Delta T n)}\left(\frac{e^{\zeta(\phi_{\zeta',n})}A_n}{2\zeta} + \frac{e^{\zeta(\phi_{\zeta'',n})}B_n}{2}\right)\zeta'}_{q_n^+} \underbrace{- e^{-\zeta(2\pi f \Delta T n)}\left(\frac{e^{-\zeta(\phi_{\zeta',n})}A_n}{2\zeta} + \frac{e^{-\zeta(\phi_{\zeta'',n})}B_n}{2}\right)\zeta'}_{q_n^-}$$

where q_n has been divided into the two counter-rotating signals q_n^+ and q_n^-, which can be expressed by the corresponding first-order quaternion linear regressions

$$q_n^+ = e^{\zeta(2\pi f \Delta T)} q_{n-1}^+ \qquad \text{and} \qquad q_n^- = e^{-\zeta(2\pi f \Delta T)} q_{n-1}^-. \tag{6.6}$$

The geometrical interpretation of system voltages and the counter-rotating elements, for an unbalanced three-phase system, is illustrated in Figure 6.2.

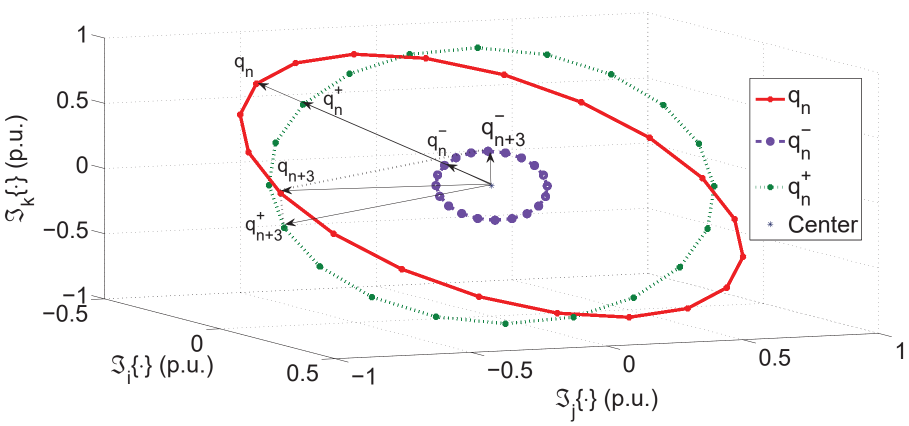

Figure 6.2: System voltage, q_n, positive sequenced element, q_n^+, and negative sequence element, q_n^-, of an unbalanced three-phase system suffering from an 80% drop in the amplitude of $v_{a,n}$ and 20 degree shifts in the phases of $v_{b,n}$ and $v_{c,n}$.

Taking into account the linear regressions in (6.6), where the phase incrementing element of q_n^+ is the quaternion conjugate of the phase incrementing element of q_n^-, a state space model for q_n is proposed in Algorithm 8, where $\varphi_n = e^{\zeta 2\pi f \Delta T}$, $\boldsymbol{\nu}_n$ is the state evolution noise, and ω_n is the observation noise. Note that the developed quaternion frequency estimator in Algorithm 8 can be implemented using the strictly linear quaternion Kalman filter. In addition, if the observation and state vectors are considered in their augmented formulation[1]; then, Algorithm 8 can be implemented using the AQKF which in turn will allow improper observational noise to be accounted for, albeit at the cost of an increase in computational complexity.

Algorithm 8. Quaternion Frequency Estimator (QFE)

State evolution equation:

$$\begin{bmatrix} \varphi_{n+1} \\ q^+_{n+1} \\ q^-_{n+1} \end{bmatrix} = \begin{bmatrix} \varphi_n \\ \varphi_n q^+_n \\ \varphi^*_n q^-_n \end{bmatrix} + \boldsymbol{\nu}_n$$

Observation equation:

$$q_n = \begin{bmatrix} 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \varphi_n \\ q^+_n \\ q^-_n \end{bmatrix} + \omega_n$$

Estimate of frequency:

$$\hat{f}_n = \frac{|\Im\left(\ln\left(\varphi_n\right)\right)|}{2\pi\Delta T} = \frac{1}{2\pi\Delta T}\text{atan}\left(\frac{|\Im(\varphi_n)|}{\Re(\varphi_n)}\right)$$

6.4 Phasor estimation

Through applying simple mathematical manipulations, the expression in (6.1) is rearranged to give

$$q_n = \Lambda_{I,n}\cos(2\pi f\Delta Tn) - \Lambda_{Q,n}\sin(2\pi f\Delta Tn) \tag{6.7}$$

where $\Lambda_{I,n}$ and $\Lambda_{Q,n}$ are given by

$$\begin{aligned} \Lambda_{I,n} =& iV_{a,n}\cos(\phi_{a,n}) + jV_{b,n}\cos(\phi_{b,n} + \frac{2\pi}{3}) + kV_{c,n}\cos(\phi_{c,n} + \frac{4\pi}{3}) \\ \Lambda_{Q,n} =& iV_{a,n}\sin(\phi_{a,n}) + jV_{b,n}\sin(\phi_{b,n} + \frac{4\pi}{3}) + kV_{c,n}\sin(\phi_{c,n} + \frac{4\pi}{3}). \end{aligned} \tag{6.8}$$

[1] In this setting, considering that the observation, q_n, has a vanishing real component we have $q_n = -q^i_n - q^j_n - q^k_n$. Therefore, q_n is not linearly independent from q^i_n, q^j_n, and q^k_n; hence, q_n need not to be included in the augmented observation vector which takes the form $q^a_n = [q^i_n, q^j_n, g^k_n]^T$. The same applies for the state vector as q^+_n and q^-_n also have vanishing real components.

Replacing $\sin(\cdot)$ and $\cos(\cdot)$ functions in the expression in (6.8) with their polar representations, gives

$$q_n = \frac{\Lambda_{I,n}}{2}\left(e^{(\zeta 2\pi f \Delta T n)} + e^{-(\zeta 2\pi f \Delta T n)}\right) + \frac{\Lambda_{Q,n}}{2\zeta}\left(e^{(\zeta 2\pi f \Delta T n)} - e^{-(\zeta 2\pi f \Delta T n)}\right) \tag{6.9}$$

where

$$\zeta = \frac{\Lambda_{I,n} \times \Lambda_{Q,n}}{|\Lambda_{I,n} \times \Lambda_{Q,n}|}.$$

From the expressions in (6.7) and (6.8), it becomes clear that $\Lambda_{I,n}$ and $\Lambda_{Q,n}$ consist of the real and imaginary components of the voltage phasors; thus, the problem of estimating the voltage phasors of the three-phase system is reduced to estimating $\Lambda_{I,n}$ and $\Lambda_{Q,n}$. In Appendix C, a deterministic relation is established between the state vector elements of the developed QFE and the voltage phasors. In addition to a deterministic relation between the state vector elements of the QFE and three-phase voltage phasors, an adaptive approach can also be taken for estimating the voltage phasors.

From the expression in (6.7), notice that $\Lambda_{I,n}$ and $\Lambda_{Q,n}$ can be estimated through demodulating the quaternion-valued signal q_n. To this end, the state space model in Algorithm 9 is proposed, where q_n^+ and q_n^-, that are elements of the state space vector of Algorithm 8, are combined together to generate the observation signal in order to prevent the observational noise, resulting from measurement inaccuracy, harmonics, and other irregularities, from affecting the phasor estimates. Note that in Algorithm 9, the observation equation is dependent on the signals $\sin(2\pi f \Delta T n)$ and $\cos(2\pi f \Delta T n)$, which are generated recursively using $\sin(2\pi f \Delta T)$ and $\cos(2\pi f \Delta T)$ that in turn are extracted from $\varphi_n = e^{\zeta 2\pi f \Delta T} = \cos(2\pi f \Delta T) + \zeta \sin(2\pi f \Delta T)$. The phasor estimation process is illustrated in Figure 6.3.

Algorithm 9. Quaternion Adaptive Phasor Estimator (QAPE)

State evolution equation:

$$\begin{bmatrix} \Lambda_{I,n} \\ \Lambda_{Q,n} \end{bmatrix} = \begin{bmatrix} \Lambda_{I,n-1} \\ \Lambda_{Q,n-1} \end{bmatrix} + \boldsymbol{\nu}_n$$

Observation equation:

$$q_n^+ + q_n^- = \begin{bmatrix} \cos(2\pi f \Delta T n) & \sin(2\pi f \Delta T n) \end{bmatrix} \begin{bmatrix} \Lambda_{I,n} \\ \Lambda_{Q,n} \end{bmatrix} + \omega_n$$

Sinusoidal Signal Generator (SSG):

$$\begin{bmatrix} \cos(2\pi f \Delta T n) \\ \sin(2\pi f \Delta T n) \end{bmatrix} = \begin{bmatrix} \Re(\varphi_n) & -|\Im(\varphi_n)| \\ |\Im(\varphi_n)| & \Re(\varphi_n) \end{bmatrix} \begin{bmatrix} \cos(2\pi f \Delta T (n-1)) \\ \sin(2\pi f \Delta T (n-1)) \end{bmatrix}$$

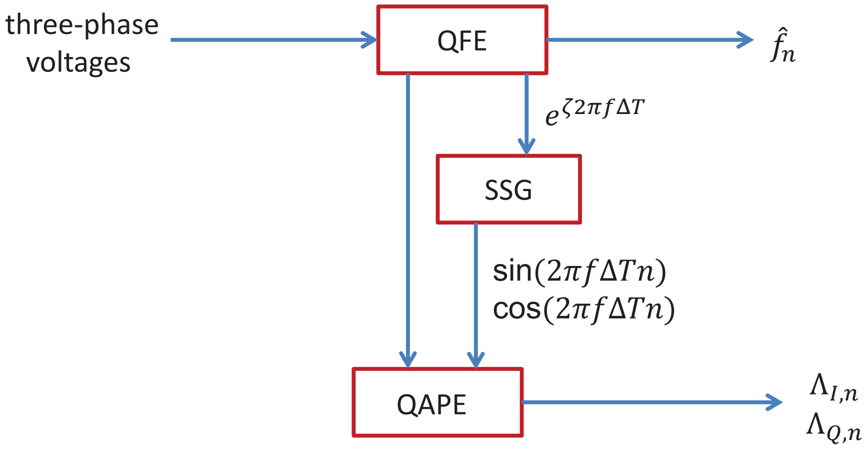

Figure 6.3: Schematic of the proposed quaternion frequency and phasor estimator. The quaternion frequency estimator (QFE) presented in Algorithm 8 and quaternion adaptive phasor estimator (QAPE) presented in Algorithm 9 with its sinusoidal signal generator (SSG) are shown.

6.5 Performance analysis

The developed QFE is implemented using the AQKF, presented in Algorithm 2, which produces unbiased estimates of the state vector and an estimate of the covariance matrix of the state vector estimation error. Therefore, the estimate of the phase incrementing element of the QFE can be modeled as

$$\hat{\varphi}_n = \varphi_n + u_n$$

where $\hat{\varphi}_n$ is the estimate of the phase incrementing element at time instant n and u_n is a zero-mean quaternion-valued Gaussian random variable representing the estimation error, the second-order statistics of which are also estimated by the AQKF. The estimate of the system frequency obtained by the QFE can now be expressed as

$$\hat{f}_n = \frac{1}{2\pi\Delta T}\text{atan}\left(\frac{|\Im(\hat{\varphi}_n)|}{\Re(\hat{\varphi}_n)}\right) = \frac{1}{2\pi\Delta T}\text{atan}\left(\frac{|\Im(\varphi_n) + \Im(u_n)|}{\Re(\varphi_n) + \Re(u_n)}\right). \tag{6.10}$$

Assuming that the sampling frequency is sufficiently high so that $\cos(2\pi f \Delta T) \approx 1$ and the variance of the noise element, u_n, is much less than one, we have

$$\Re(\varphi_n) + \Re(u_n) = \cos(2\pi f \Delta T) + \Re(u_n) \approx \cos(2\pi f \Delta T). \tag{6.11}$$

Furthermore, considering that $\Im(\varphi_n) = \zeta \sin(2\pi f \Delta T)$, the expression $|\Im(\varphi_n) + \Im(u_n)|$ in (6.10) gives

$$\begin{aligned} |\Im(\varphi_n) + \Im(u_n)| &= \sqrt{\left(\Im(\varphi_n) + \Im(u_n)\right)\left(\Im(\varphi_n^*) + \Im(u_n^*)\right)} \\ &= \sqrt{\sin^2(2\pi f \Delta T) + |\Im(u_n)|^2 + 2\Re\left(\zeta\Im(u_n)\right)\sin(2\pi f \Delta T)} \end{aligned} \tag{6.12}$$

where the square root function in (6.12) can be approximated by its first-order Taylor expansion around $\sin^2(2\pi f \Delta T)$ to give

$$|\Im(\varphi_n) + \Im(u_n)| = \sin(2\pi f \Delta T) + \Re(\zeta \Im(u_n)) + \frac{|\Im(u_n)|^2}{2\sin(2\pi f \Delta T)}. \tag{6.13}$$

Now, replacing (6.13) and (6.11) into the expression in (6.10) yields

$$\hat{f}_n = \frac{1}{2\pi \Delta T}\text{atan}\left(\frac{\sin(2\pi f \Delta T)}{\cos(2\pi f \Delta T)} + \frac{\Re(\zeta \Im(u_n))}{\cos(2\pi f \Delta T)} + \frac{|\Im(u_n)|^2}{2\cos(2\pi f \Delta T)\sin(2\pi f \Delta T)}\right). \tag{6.14}$$

Again assuming the sampling frequency is sufficiently high and the variance of the noise element, u_n, is much less than one, allows the atan$(\cdot)$ functon in (6.14) to be approximated by its first-order Taylor expansion to give

$$\hat{f}_n = f + \frac{\Re(\zeta \Im(u_n))}{2\pi \Delta T} + \frac{|\Im(u_n)|^2}{8f(\pi \Delta T)^2}. \tag{6.15}$$

Taking the statistical expectation of the expression in (6.15) yields

$$E\left[\hat{f}_n\right] = f + \frac{C_{\Im(u_n)}}{8f(\pi \Delta T)^2}. \tag{6.16}$$

In addition, after some tedious mathematical manipulations the MSE of the frequency estimates obtained by the QFE can be expressed as

$$E\left[\left(\hat{f}_n - f\right)^2\right] = \frac{E\left[|\Re(\zeta \Im(u_n))|^2\right]}{4(\pi \Delta T)^2} + \frac{E\left[|\Im(u_n)|^4\right]}{64f^2(\pi \Delta T)^4} + \frac{E\left[\Re(\zeta \Im(u_n))|\Im(u_n)|^2\right]}{8f(\pi \Delta T)^3}. \tag{6.17}$$

From the inequality $\Re(\zeta \Im(u_n)) \le |\Im(u_n)|^2$, we have

$$E\left[|\Re(\zeta \Im(u_n))|^2\right] \le E\left[|\Im(u_n))|^4\right] \quad \text{and} \quad E\left[\Re(\zeta \Im(u_n))|\Im(u_n)|^2\right] \le E\left[|\Im(u_n))|^4\right]. \tag{6.18}$$

Now, replacing the inequalities in (6.18) into the expression (6.17) allows an upper bound to be established for the MSE of the frequency estimates obtained by the QFE that is given by

$$E\left[\left(\hat{f}_n - f\right)^2\right] \le \kappa \left(\frac{1}{4(\pi \Delta T)^2} + \frac{1}{64f^2(\pi \Delta T)^4} + \frac{1}{8f(\pi \Delta T)^3}\right) \tag{6.19}$$

where $\kappa = E\left[|\Im(u_n))|^4\right]$ which, using the framework established in Appendix A, can be expressed as

$$\kappa = E\left[|\Im(u_n))|^4\right] = \frac{3|C_{\Im(u_n)}|^2 + |R_{\Im(u_n)\Im^i(u_n)}|^2 + |R_{\Im(u_n)\Im^j(u_n)}|^2 + |R_{\Im(u_n)\Im^k(u_n)}|^2}{2}.$$

The expressions in (6.16), (6.17), and (6.19) show the effect of operating conditions, such as observational noise and system frequency, on the bias and MSE performance of the developed QFE. The performance of the QFE is compared to that of its complex-valued counterparts in a more practical setting in Section 6.7 using simulations on synthetically generated data and real-world data recordings.

6.6 Harmonic contamination

In Section 5.5, two methods for dealing with harmonic contamination were presented, accounting harmonics as observational noise and dealing with harmonics as part of the signal. However, as it was mentioned in Section 5.5 and indicated by the analysis in Section 6.5, accounting harmonics as observational noise, leads to a performance loss in both bias and MSE performance. Although the harmonic components can be incorporated into the state space model of the QFE, this approach leads to a computationally complex algorithm that requires the inversion of relatively large quaternion-valued matrices. Theretofore, in this section a method for adaptive cancellation of the major harmonic components of the three-phase power signal is presented.

Taking the same approach as in Section 6.3, the m^{th} harmonic component of the three-phase power signal, denoted by $q_{n,m}$, can also be divided into two counter-rotating elements so that $q_{n,m} = q^{+}_{n,m} + q^{-}_{n,m}$ with

$$q^{+}_{n,m} = e^{\zeta_m (2m\pi f \Delta T)} q^{+}_{n-1,m} \qquad \text{and} \qquad q^{-}_{n,m} = e^{-\zeta_m (2m\pi f \Delta T)} q^{-}_{n-1,m} \tag{6.20}$$

where ζ_m is a unit quaternion designed to be normal to the plane that contains $q_{n,m}$ and $\zeta_m^2 = -1$. Therefore, the evolution of the m^{th} harmonic component of the three-phase power signal can be modeled as

$$\underbrace{\begin{bmatrix} \varphi_{n+1,m} \\ q^{+}_{n+1,m} \\ q^{-}_{n+1,m} \end{bmatrix}}_{\mathbf{x}_{n+1,m}} = \underbrace{\begin{bmatrix} \varphi_{n,m} \\ \varphi_{n,m} q^{+}_{n,m} \\ \varphi^{*}_{n,m} q^{-}_{n,m} \end{bmatrix}}_{f_{n,m}(\mathbf{x}_{n,m})} + \boldsymbol{\nu}_{n,m} \tag{6.21}$$

where $\mathbf{x}_{n,m} = [\varphi_{n+1,m}, q^{+}_{n+1,m}, q^{-}_{n+1,m}]^T$ and $f_{n,m}(\cdot)$ are the state vector and state evolution function of the m^{th} harmonic component at time instant n, while $\boldsymbol{\nu}_{n,m}$ denotes the state evolution noise. In this setting, if the fundamental system frequency, f, is known, the m^{th} harmonic component of the three-phase power signal can be tracked using the constrained Kalman filter in Algorithm 10, where the transform

$$\underbrace{\begin{bmatrix} e^{\zeta_m 2m\pi f \Delta T} \\ q^{+}_{n+1,m} \\ q^{-}_{n+1,m} \end{bmatrix}}_{\mathbf{x}'_{n,m}} = \underbrace{\begin{bmatrix} \kappa & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{T}_{n,m}} \underbrace{\begin{bmatrix} \varphi_{n,m} \\ q^{+}_{n+1,m} \\ q^{-}_{n+1,m} \end{bmatrix}}_{\mathbf{x}_{n,m}} \tag{6.22}$$

with

$$\kappa = \frac{e^{\zeta_m 2m\pi f \Delta T}}{\varphi_{n,m}} \qquad \text{and} \qquad \zeta_m = \frac{\Im(\varphi_{n,m})}{|\Im(\varphi_{n,m})|}$$

is used to constrain the state vector estimates to those with phase incrementing elements at m times the fundamental frequency of the system, while $\mathbf{A}^a_{n,m}$ is the Jacobian matrix of $f_{n,m}$ and $\hat{\mathbf{q}}^a_n = \sum_{\forall m} \mathbf{H}^a \hat{\mathbf{x}}^{a'}_{n,n-1,m}$ gives the estimate of the augmented observation vector, $\mathbf{q}^a_n$, that accounts for the main component of the signal, that is $m = 1$, and harmonic components for which $\hat{\mathbf{x}}^{a'}_{n,n-1}$ is available[2].

Algorithm 10. Quaternion Harmonic Estimator (QHE)

Initialize with:

$$\begin{aligned}
\hat{\mathbf{x}}^{a'}_{0|0,m} &= E[\mathbf{x}^{a'}_{0,m}] \\
\hat{\mathbf{M}}^{a'}_{0|0,m} &= E\big[(\mathbf{x}^{a'}_{0,m} - E[\mathbf{x'}^{a}_{0,m}])(\mathbf{x'}^{a}_{0,m} - E[\mathbf{x}^{a'}_{0,m}])^H\big]
\end{aligned}$$

Model update:

$$\begin{aligned}
\hat{\mathbf{x}}^{a}_{n|n-1,m} &= f_{n,m}(\hat{\mathbf{x}}^{a'}_{n-1|n-1,m}) \\
\hat{\mathbf{M}}^{a}_{n|n-1,m} &= \mathbf{A}^a_{n,m}\hat{\mathbf{M}}^{a'}_{n-1|n-1,m}\mathbf{A}^{aH}_{n,m} + \mathbf{C}_{\boldsymbol{\nu}^a_{n,m}}
\end{aligned}$$

Measurement update:

$$\begin{aligned}
\hat{\mathbf{M}}^{a^{-1}}_{n|n,m} &= \hat{\mathbf{M}}^{a^{-1}}_{n|n-1,m} + \mathbf{H}^{aH}\mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{n,m}}\mathbf{H}^a \\
\mathbf{G}^a_{n,m} &= \hat{\mathbf{M}}^{a}_{n|n,m}\mathbf{H}^{aH}\mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{n,m}} \\
\hat{\mathbf{x}}^{a}_{n|n,m} &= \hat{\mathbf{x}}^{a}_{n|n-1,m} + \mathbf{G}^a_{n,m}\left(\mathbf{q}^a_n - \sum_{\forall m}\mathbf{H}^a\hat{\mathbf{x}}^a_{n|n-1,m}\right)
\end{aligned}$$

Implement constraint transform as given in (6.22):

$$\hat{\mathbf{x}}^{a'}_{n|n,m} = \mathbf{T}^a_{n,m}\hat{\mathbf{x}}^a_{n|n,m}$$

Update the augmented covariance matrix of the error:

$$\hat{\mathbf{M}}^{a'}_{n|n,m} = \mathbf{T}^a_{n,m}\hat{\mathbf{M}}^a_{n|n,m}\mathbf{T}^{aH}_{n,m}$$

Now, the harmonic components of the three-phase power signal can be dealt with through cooperation between a number of AQKFs, one of which estimates the fundamental frequency of the system whereas the other AQKFs use the estimates of the fundamental system frequency to calculate the m^{th} harmonic component of the signal, which is in turn utilized to cancel the effect of the m^{th} harmonic component in the observation of the AQKF that is estimating the fundamental system frequency of the three-phase power system. In Figure 6.4 a schematic of the proposed method is

[2] It is important to note that if estimates of the state vector of the m^{th} harmonic component of the three-phase power signal is at hand; then, its voltage phasors can also be obtained through implementing the QAPE developed in Section 6.4.

shown, where the modified quaternion frequency estimator (MQFE) estimates the main frequency of the systems whereas the quaternion harmonic estimator (QHE) estimates the m^{th} harmonic component of the signal. In addition, a schematic illustrating the operations of the MQFE is shown in Figure 6.5, where the harmonic components of the signal for which estimates of their state vectors, $\hat{\mathbf{x}}^{a'}_{n|n,m}$, are obtained through the QHE are adaptively canceled in the observed three-phase power signal by deducting the term '$\sum_{\forall m} \mathbf{H}^a \hat{\mathbf{x}}^{a'}_{n|n-1,m}$' from the observational signal before implanting the QFE that was developed in Section 6.3 for adaptive frequency estimation.

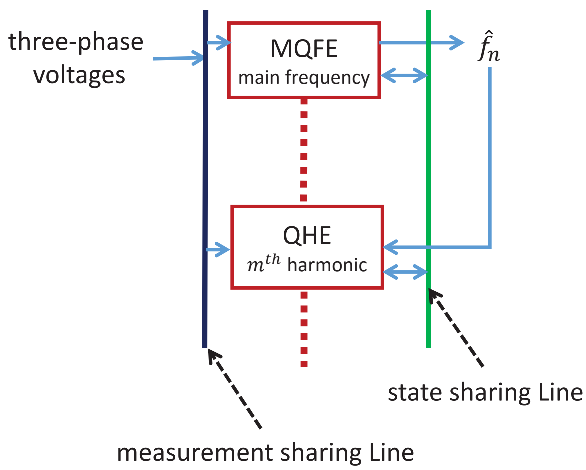

Figure 6.4: Schematic of the proposed technique for dealing with harmonics showing the QFE that estimates the main frequency of the system and the QHE that estimate the m^{th} harmonic component of the signal and its contribution.

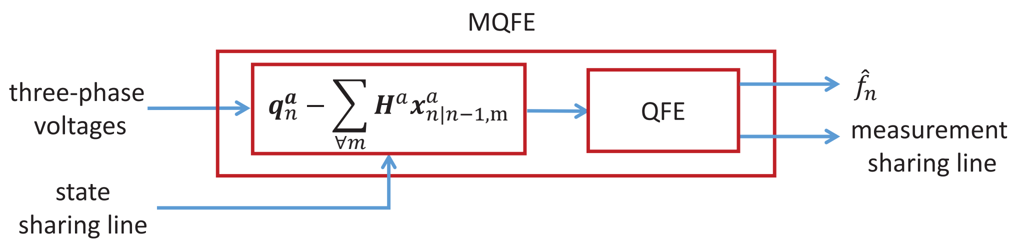

Figure 6.5: Schematic of the operations of the modified quaternion frequency estimator (MQFE). The harmonic components of the signal for which estimates of their state vectors, $\hat{\mathbf{x}}^{a'}_{n|n,m}$, are available through the QHEs are used to adaptively cancel their effect in the observational signal by deducting the term '$\sum_{\forall m} \mathbf{H}^a \hat{\mathbf{x}}^{a'}_{n|n-1,m}$' from the observational signal before implementing the QFE for frequency estimation.

6.7 Simulations

In this section, the performance of the developed QFE is assessed and benchmarked against those of its strictly linear and widely-linear complex-valued counterparts in a number of experiments involving practical power grid scenarios using both synthetically generated data and real-world data recordings. In all experiments the sampling frequency was $f_s = 1$ kHz and the voltage measurements were considered to be corrupted by white Gaussian noise with SNR of 40 dB.

6.7.1 Synthetically generated data

In the first experiment, the three-phase system was considered to be initially operating at its nominal frequency of 50 Hz and in a balanced fashion, then the system suffered a fault characterized by an 80% drop in the amplitude of $v_{a,n}$ and 20 degree shifts in the phases of $v_{b,n}$ and $v_{c,n}$; in addition, the frequency of the system experienced a step jump of 2 Hz. The fault lasted for a short duration and the system returned to balanced operating conditions and its nominal frequency once more. The estimates of the system frequency obtained through implementing the CNLFE, CWLFE and QFE are shown in Figure 6.6. Observe that all of the algorithms could accurately estimate the system frequency with the QFE achieving a lower steady-state variance than its complex-valued counterparts. Moreover, a closer examination of the transient behavior of the CNLFE, CWLFE and QFE is made in Figure 6.7. Notice that both when the three-phase system was starting to experience the fault and was recovering from the fault, the QFE outperformed the CWLFE in terms of convergence rate. Furthermore, the estimates of the system voltage phasors of the three-phase power system obtained trough implementing the QAPE, developed in Section 6.4, are shown in Figure 6.8. Notice that the developed algorithm accurately tracks the voltage phasors of the system.

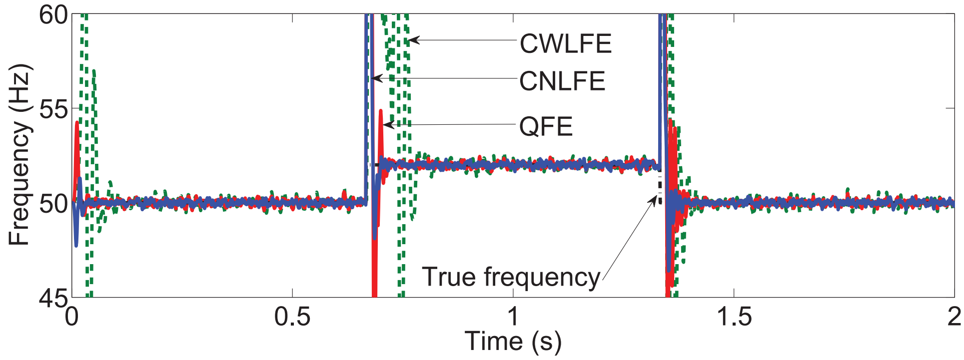

Figure 6.6: Frequency estimation for a three-phase power system experiencing a short fault and a 2 Hz jump in frequency from 0.667 to 1.334 seconds. The performance of the QFE is shown along side those obtained by implementing CWLFE and CNLFE.

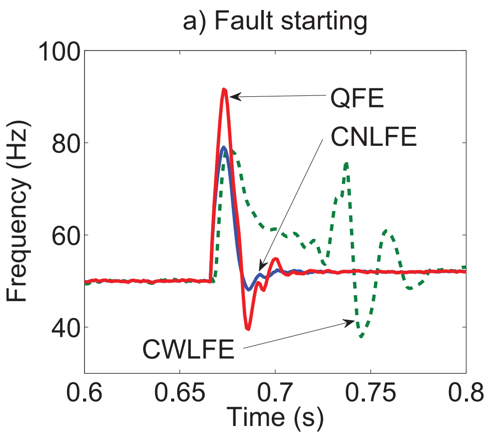

Figure 6.7: Transient behavior of the QFE, CNLFE and CWLFE when the three-phase power system experiences a fault and a 2 Hz step jump in frequency: a) fault starting, b) fault ending.

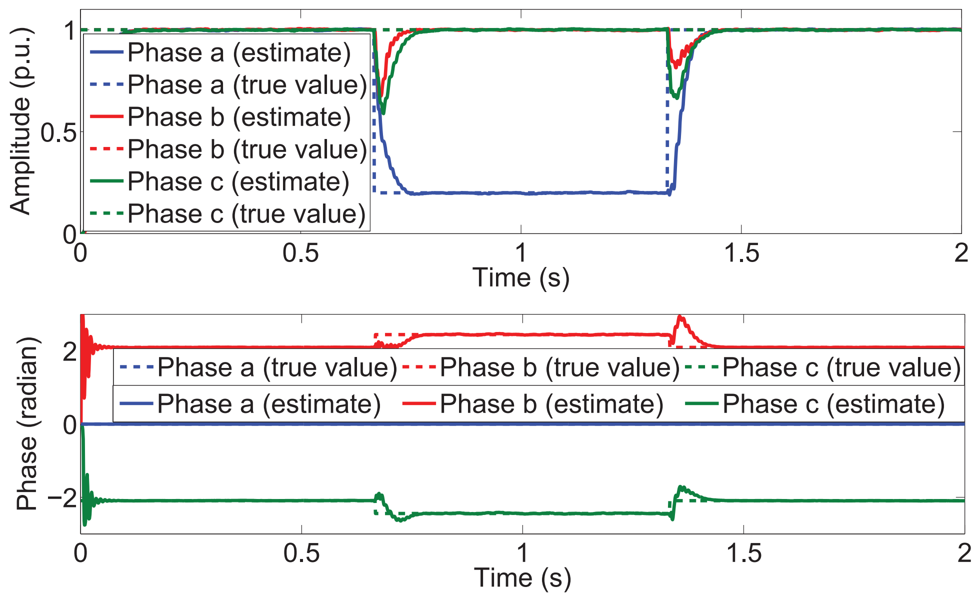

Figure 6.8: Voltage phasor estimation for a three-phase system experiencing a short voltage sag from 0.667 to 1.334 seconds implementing the QAPE. The amplitude of the voltage phasors are shown in the top graph and the phase angles of the voltage phasors are shown in the bottom graph.

In the second experiment, a three-phase system operating under unbalanced conditions caused by an 80% drop in the amplitude of $v_{a,n}$ and a 20 degree shift in the phases of $v_{b,n}$ and $v_{c,n}$ (as shown in Figure 6.2), which experiences a rising (*cf.* falling) frequency due to mismatch between power generation and consumption 0.5 seconds after the simulation starts was considered. Figure 6.9 shows the estimates of the system frequency obtained through implementing the developed QFE. Observe that the QFE accurately tracked the system frequency phasors both when it was constant and when it was changing. In addition, the estimates of the system voltage phasors are shown in

Figure 6.10. Note that the proposed QAPE was able to accurately track the voltage phasors of the three-phase power system both when the system frequency was constant and when the system frequency was changing.

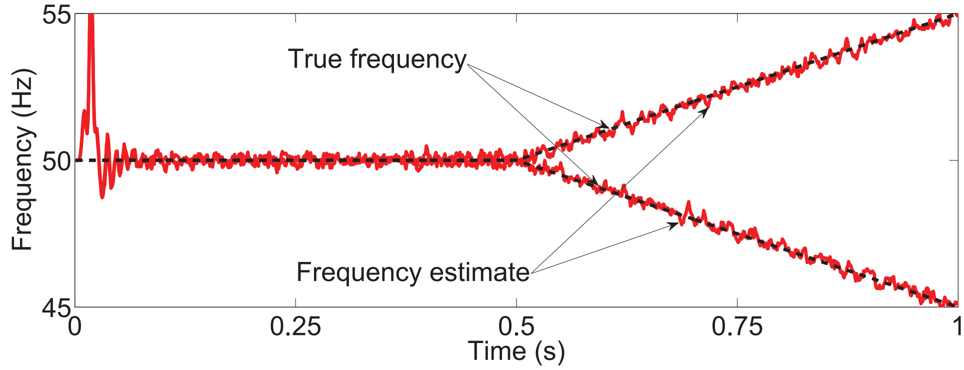

Figure 6.9: Frequency estimation for an unbalanced three-phase power system experiencing changing frequency at the rate of 10 Hz/s employing the QFE.

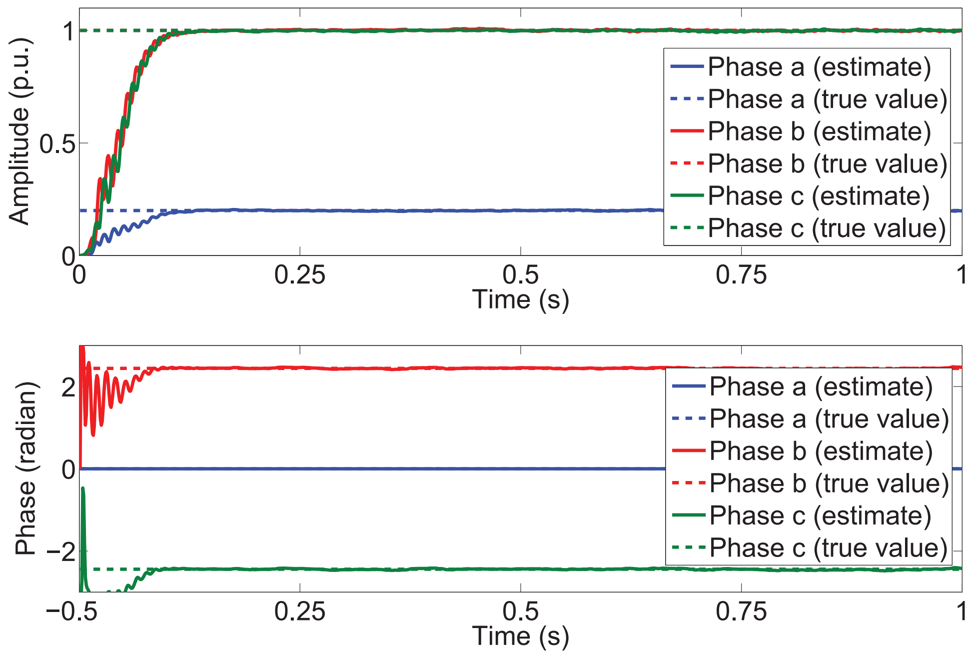

Figure 6.10: Voltage phasor estimation for an unbalanced three-phase system with changing frequency through implementing the newly developed QAPE. The amplitude of the voltage phasors are shown in the top graph and the phase angles of the voltage phasors are shown in the bottom graph.

In the third experiment, the effect of harmonics on the performance of the QFE was considered. This was performed through the addition of a balanced 10% third harmonic component to the unbalanced three-phase system used in the previous experiment at $t = 0.5$ s. The estimates of the system frequency obtained through the QFE using the framework developed in Section 6.6 are shown in Figure 6.11. In addition, voltage phasors estimates of the main frequency component of the three-phase power system obtained through the QAPE are shown in Figure 6.11. Notice that the proposed algorithm accurately estimated the system frequency and its voltage phasors regardless of the presence of the harmonic component.

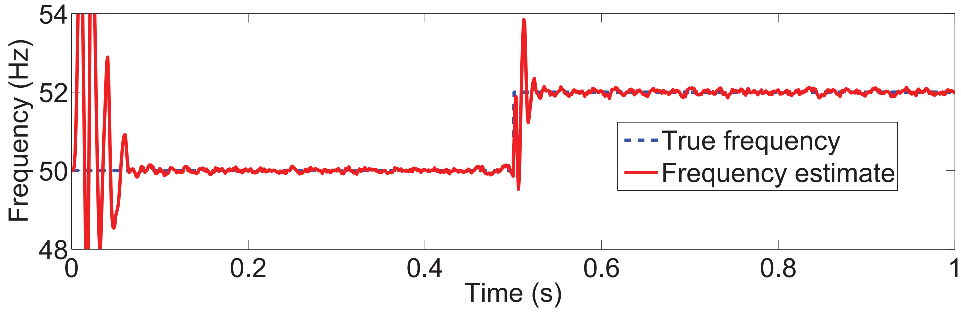

Figure 6.11: Frequency estimation for a three-phase system operating under unbalance conditions in addition to suffering from a 10% third harmonic from $t = 0.5$ s onwards.

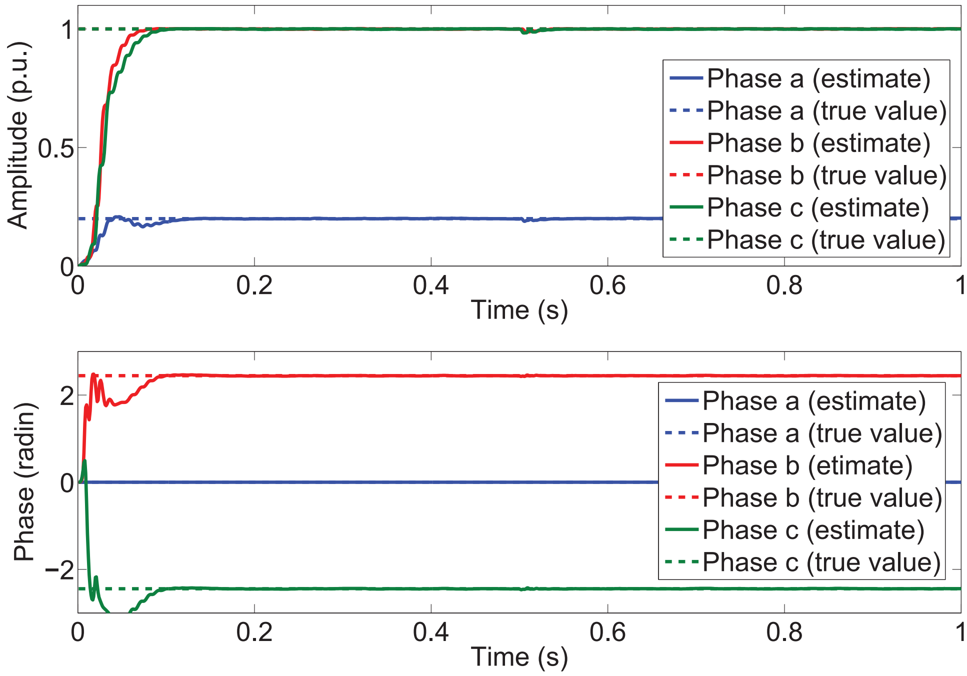

Figure 6.12: Voltage phasor estimation for a three-phase system operating under unbalanced conditions in addition to suffering from a 10% third harmonic from $t = 0.5$ s onwards. The amplitude of the voltage phasors are shown in the top graph and the phase angles of the voltage phasors are shown in the bottom graph.

The steady-state MSE performance of the newly developed QFE, for various SNR values and for both a balanced and the unbalanced three-phase system shown in Figure 6.2, is shown alongside the MSE performance of the CLFE, CWLFE, and CNLFE in Figure 6.13. Notice that the QFE achieves a smaller steady-state MSE than that of its complex-valued counterparts.

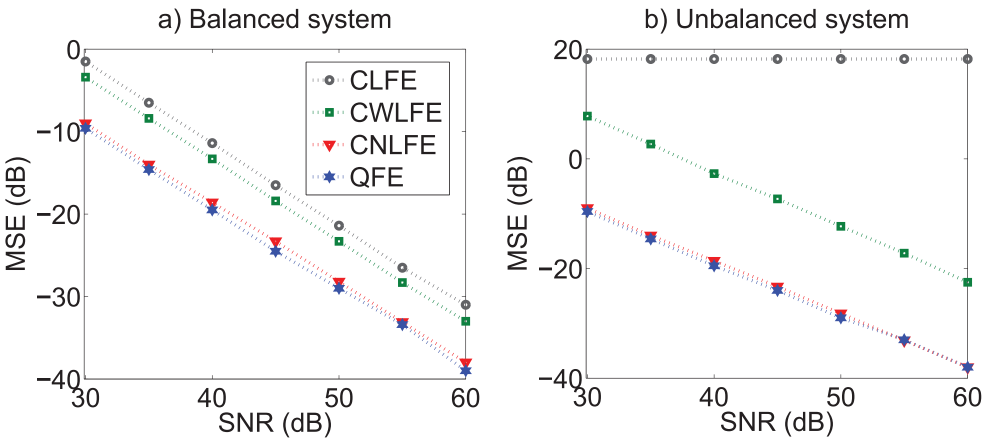

Figure 6.13: Steady-state MSE performance analysis for SNR ranging from 30 dB to 60 dB: a) balanced three-phase system, b) unbalanced three-phase system.

The steady-state bias performance of the newly developed QFE for various SNR values is now investigated. In Figure 6.14, the steady-state bias performance of the QFE is compared to that of the CNLFE, CWLFE and the CLFE for both a balanced three-phase system and the unbalanced three-phase system characterized in Figure 6.2. Observe that the developed QFE achieves a lower bias than the CWLFE and CLFE.

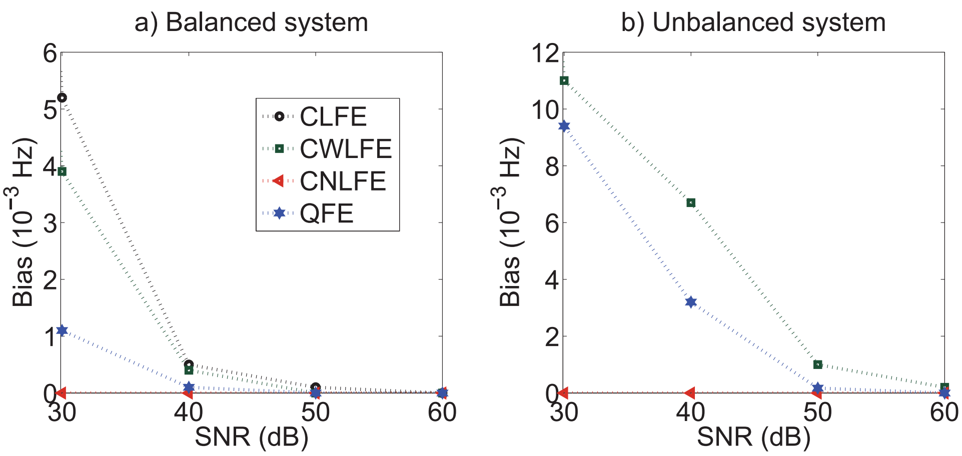

Figure 6.14: Steady-state bias performance analysis for SNR ranging from 30 dB to 60 dB: a) balanced three-phase system, b) unbalanced three-phase system.

6.7.2 Real-world data

The performance of the developed QFE is now assessed during a voltage sag using real-world data. The recorded data and the performance of the QFE is shown in Figure 6.15, where the frequency estimates obtained through the CWLFE are provided for comparison. Notice the QFE showed outstanding performance both during nominal operating conditions and during the voltage sag. Furthermore, The performance of the proposed

QAPE is shown in Figure 6.16, where the root mean square (RMS) of the voltage phasors are displayed as base-line truth.

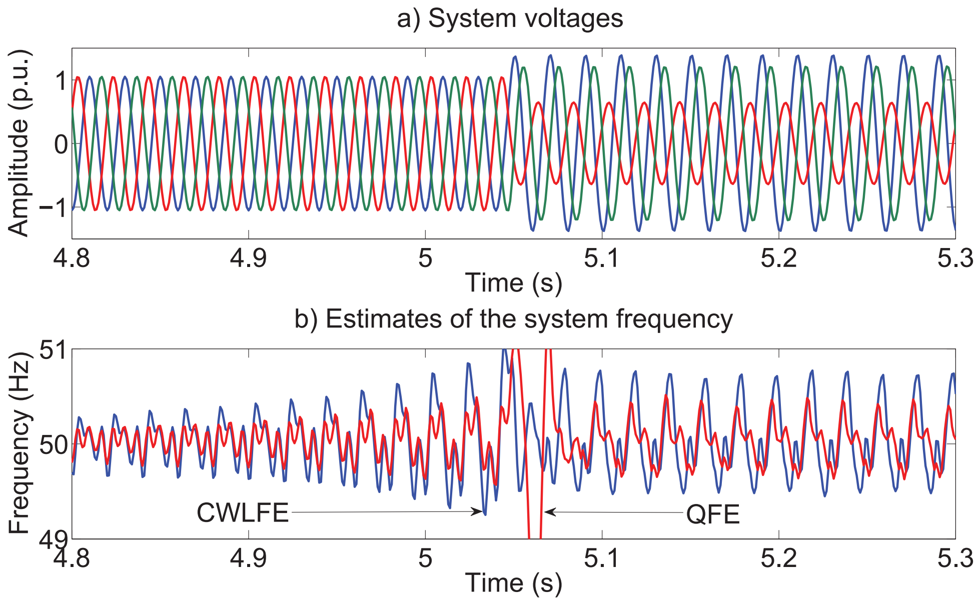

Figure 6.15: Frequency estimation using real-world data recording during a voltage sag: a) system voltages, b) estimates of the system frequency obtained through implementing the QFE and CWLFE.

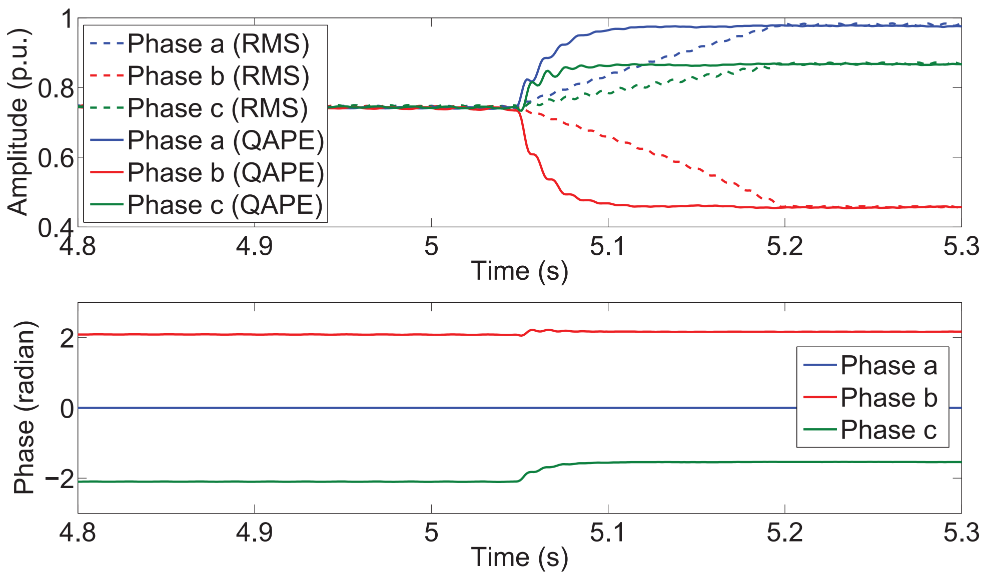

Figure 6.16: Voltage phasor estimation using real-world data recording during a voltage sag implementing the QAPE. The amplitude of the voltage phasors are shown in the top graph and the phase angles of the voltage phasors are shown in the bottom graph.

6.8 Conclusion

A novel algorithm has been developed for joint estimation of the fundamental frequency and voltage phasors of three-phase power systems. The proposed algorithm exploits the multidimensional nature of quaternions to make possible the full characterization of three-phase power systems in the three-dimensional domain where they naturally reside and eliminates the need for using the Clarke transform, allowing the components of the designed quaternion-valued state space vector to be used for joint adaptive estimation of the system frequency and voltage phasors. The performance of the proposed algorithm has been assessed in a number of scenarios using both synthetically generated and real-world data, where it has shown outstanding performance and outperformed its complex-valued counterparts.

Chapter 7

A Distributed Quaternion Kalman Filter

7.1 Overview

Most of the research surrounding adaptive distributed signal processing is conducted in the real and complex domains, whereas in many real-world applications the data sources are three-dimensional by nature, offering an opportunity for quaternions in terms of convenience of representation and mathematical tractability. In this chapter, we expand the concept of distributed Kalman filtering to the quaternion domain in order to develop a robust distributed quaternion Kalman filtering algorithm for data fusion over sensor networks dealing with three-dimensional data. For rigor, the mean and mean square behavior of the algorithm is analyzed. Finally, the developed algorithm is used to estimate the main system frequency in power distribution networks and for collaborative target tracking.

7.2 Introduction

In recent years, sensor networks have been used in a variety of applications, such as collaborative target tracking, distributed fault detection, control of unmanned aerial vehicles, and automated vehicle guidance technology [124–134]. In these applications, algorithms based on Kalman filtering have proven to be advantageous offering better accuracy and faster convergence rates due to their underlying state space model that accounts for observational noise. In addition, owing to the low implementation cost and computational efficiency that distributed estimation and tracking techniques offer, as compared to their centralized counterparts, distributed signal processing algorithms, have proven to be suitable for real-time implementation, computationally efficient, scalable with the size of the network, and robust to link failure [130–132].

In light of the advantages that quaternion-valued signal processing algorithms offer, we expand the framework of quaternion-valued Kalman filtering to the distributed setting in order to develop a truly distributed quaternion Kalman filter applicable for frequency estimation in three-phase power distribution networks and collaborative target tracking, where quaternions offer the dimensionality necessary to model such signals directly in the multi-dimensional domain that they originate from. The distributed quaternion Kalman filter is developed through decomposing the operations of the centralized quaternion Kalman filter in such a fashion that they can be performed locally by the individual nodes (sensors) of the network.

The performance analysis of the developed algorithm shows that it operates in an unbiased fashion. Moreover, in order to establish the effects of correlated observation noise in the network on the performance of the developed algorithm and quantify its mean square performance, a recursive expression for the augmented covariance matrix of the estimation error at each node is derived. The performance of the derived distributed quaternion Kalman filter is illustrated in smart grid applications for estimating the fundamental frequency of three-phase power distribution networks and for collaborative target tracking in a bearings-only scenario, where each sensor in the network can only observe the bearings of the target.

7.3 The distributed quaternion Kalman filter

Consider a set of sensors denoted by $\mathcal{N}$ that are interconnected in a network and let the neighborhood of a node to be the subset of nodes that communicate with the node, including self-communication. Organizing all observations made by different nodes throughout the network in the column vector

$$\mathbf{y}^a_{col,n} = [\mathbf{y}^{aT}_{1,n}, \dots, \mathbf{y}^{aT}_{|\mathcal{N}|,n}]^T$$

where $\mathbf{y}^a_{m,n}$ represents the augmented observation vector at node m at time n and $|\mathcal{N}|$ denotes the number of nodes in the network, allows the augmented state vector sequence to be estimated by the centralized augmented quaternion Kalman filter (CAQKF) given in Algorithm 11, where

$$\mathbf{H}^a_{col,n} = [\mathbf{H}^{aT}_{1,n}, \dots, \mathbf{H}^{aT}_{|\mathcal{N}|,n}]^T$$

is the column block matrix of the augmented observation functions with $\mathbf{H}^a_{m,n}$ representing the observation function at node m and at time instant n, while $\mathbf{C}_{\boldsymbol{\omega}^a_{col,n}}$ is the augmented covariance matrix of the column vector of the combined augmented observational noises given by

$$\boldsymbol{\omega}^a_{col,n} = [\boldsymbol{\omega}^{aT}_{1,n}, \ldots, \boldsymbol{\omega}^{aT}_{|\mathcal{N}|,n}]^T$$

with $\boldsymbol{\omega}^a_{m,n}$ denoting the observational noise at node m and at time instant n, whereas $\mathbf{A}^a_n$ represents the augmented state evolution function time instant n.

Algorithm 11 . Centralized Augmented Quaternion Kalman Filter (CAQKF)

Initialize with:

$$\begin{aligned}\hat{\mathbf{x}}^a_{0|0} &= E[\mathbf{x}^a_0] \\ \hat{\mathbf{M}}^a_{0|0} &= E\big[(\mathbf{x}^a_0 - E[\mathbf{x}^a_0])(\mathbf{x}^a_0 - E[\mathbf{x}^a_0])^H\big]\end{aligned}$$

Model update:

$$\begin{aligned}\hat{\mathbf{x}}^a_{n|n-1} &= \mathbf{A}^a_n \hat{\mathbf{x}}^a_{n-1|n-1} \\ \hat{\mathbf{M}}^a_{n|n-1} &= \mathbf{A}^a_n \hat{\mathbf{M}}^a_{n-1|n-1} \mathbf{A}^{aH}_n + \mathbf{C}_{\boldsymbol{\nu}^a_n}\end{aligned}$$

Measurement update:

$$\begin{aligned}\hat{\mathbf{M}}^{a^{-1}}_{n|n} &= \hat{\mathbf{M}}^{a^{-1}}_{n|n-1} + \mathbf{H}^{aH}_{col,n} \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{col,n}} \mathbf{H}^a_{col,n} \\ \mathbf{G}^a_n &= \hat{\mathbf{M}}^a_{n|n} \mathbf{H}^{aH}_{col,n} \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{col,n}} \\ \hat{\mathbf{x}}^a_{n|n} &= \hat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_n \big(\mathbf{y}^a_{col,n} - \mathbf{H}^a_{col,n} \hat{\mathbf{x}}^a_{n|n-1}\big)\end{aligned}$$

Although the CAQKF is optimal in the sense that it incorporates all the available information in the network, its operation requires inversions of large matrices and transfer of all observation vectors to the central node, which burdens the central node with communication traffic and heavy computations. However, assuming that the observational noise at one node is uncorrelated with the observational noise at other nodes in the network, leads to a block diagonal $\mathbf{C}_{\boldsymbol{\omega}^a_{col,n}}$ and therefore the *a posteriori* augmented state vector estimate can be expressed as

$$\hat{\mathbf{x}}^a_{n|n} = \hat{\mathbf{x}}^a_{n|n-1} + \sum_{\forall l \in \mathcal{N}} \hat{\mathbf{M}}^a_{n|n} \mathbf{H}^{aH}_{l,n} \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{l,n}} \left(\mathbf{y}^a_{l,n} - \mathbf{H}^a_{l,n} \hat{\mathbf{x}}^a_{n|n-1}\right). \tag{7.1}$$

Furthermore, the *a posteriori* augmented state vector estimate in (7.1) can be alternatively calculated by the summation

$$\hat{\mathbf{x}}^a_{n|n} = \frac{1}{|\mathcal{N}|} \sum_{\forall l \in \mathcal{N}} \boldsymbol{\phi}^a_{l,n} \tag{7.2}$$

where

$$\boldsymbol{\phi}^a_{l,n} = \hat{\mathbf{x}}^a_{n|n-1} + |\mathcal{N}| \hat{\mathbf{M}}^a_{n|n} \mathbf{H}^{aH}_{l,n} \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{l,n}} \left(\mathbf{y}^a_{l,n} - \mathbf{H}^a_{l,n} \hat{\mathbf{x}}^a_{n|n-1}\right). \tag{7.3}$$

From the CAQKF, in the formulation presented in Algorithm 11, assuming uncorrelated observation noise throughout the network we have

$$\hat{\mathbf{M}}^{a^{-1}}_{n|n} = \hat{\mathbf{M}}^{a^{-1}}_{n|n-1} + \sum_{\forall l \in \mathcal{N}} \mathbf{H}^{aH}_{l,n} \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{l,n}} \mathbf{H}^a_{l,n}. \tag{7.4}$$

Now, substituting (7.4) into (7.3) yields

$$\boldsymbol{\phi}^a_{l,n} = \hat{\mathbf{x}}^a_{n|n-1} + \mathbf{G}^a_{l,n} \left(\mathbf{y}^a_{l,n} - \mathbf{H}^a_{l,n} \hat{\mathbf{x}}^a_{n|n-1} \right) \tag{7.5}$$

where $\mathbf{G}^a_{l,n}$ is given by

$$\mathbf{G}^a_{l,n} = \left(\hat{\mathbf{M}}^{a^{-1}}_{n|n-1} + \sum_{\forall m \in \mathcal{N}} \mathbf{H}^{aH}_{m,n} \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{m,n}} \mathbf{H}^a_{m,n} \right)^{-1} \mathbf{H}^{aH}_{l,n} |\mathcal{N}| \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{l,n}}. \tag{7.6}$$

Note that, with the assumption that the network is connected[1], $\mathbf{G}^a_{l,n}$ in the formulation given in (7.6) can be obtained in a distributed fashion through the summation of local parameters $\mathbf{H}^{aH}_{m,n} \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{m,n}} \mathbf{H}^a_{m,n}$; in addition, taking into account that $\hat{\mathbf{x}}^a_{n|n}$ can be obtained by averaging local estimates, $\boldsymbol{\phi}^a_{l,n}$, allows the operations of the CAQKF to be approximated in a distributed fashion through implementing the distributed augmented quaternion Kalman filter (DAQKF) in Algorithm 12, where $\mathcal{N}_l$ denotes the set of nodes in the neighborhood of node l. The DAQKF implemented at a node is optimal in the neighborhood of that node in the scene that it operates akin to a CAQKF that combines all the information available to the nodes in its neighborhood.

7.4 Performance analysis

In order to analyze the mean and mean square performance of the developed algorithm, in this section, the error of the augmented state vector estimates is expressed in a recursive manner. The difference between the true augmented state vector and the local estimate at node l and at time instant n is given by $\boldsymbol{\epsilon}^a_{l,n} = \mathbf{x}^a_n - \boldsymbol{\phi}^a_{l,n}$ which can alternatively be expressed as

$$\boldsymbol{\epsilon}^a_{l,n} = \mathbf{x}^a_n - \hat{\mathbf{x}}^a_{l,n|n-1} - \mathbf{G}^a_{l,n} \left(\mathbf{y}^a_{l,n} - \mathbf{H}^a_{l,n} \hat{\mathbf{x}}^a_{l,n|n-1} \right). \tag{7.7}$$

Replacing $\mathbf{y}^a_{l,n} = \mathbf{H}^a_{l,n} \mathbf{x}^a_n + \boldsymbol{\omega}^a_{l,n}$ and $\boldsymbol{\epsilon}^a_{l,n|n-1} = \mathbf{x}^a_n - \hat{\mathbf{x}}^a_{l,n|n-1}$ into the expression in (7.7) yields

$$\boldsymbol{\epsilon}^a_{l,n} = \left(\mathbf{I} - \mathbf{G}^a_{l,n} \mathbf{H}^a_{l,n} \right) \boldsymbol{\epsilon}^a_{l,n|n-1} - \mathbf{G}^a_{l,n} \boldsymbol{\omega}^a_{l,n}. \tag{7.8}$$

[1]A network is referred to as connected if there exists a path between any two given nodes in the network.

Algorithm 12 . Distributed Augmented Quaternion Kalman Filter (DAQKF)

For node $l = \{1, \cdots, |\mathcal{N}|\}$:
Initialize with:

$$\begin{aligned}\hat{\mathbf{x}}^a_{l,0|0} &= E[\mathbf{x}^a_0] \\ \hat{\mathbf{M}}^a_{l,0|0} &= E\big[(\mathbf{x}^a_0 - E[\mathbf{x}^a_0])(\mathbf{x}^a_0 - E[\mathbf{x}^a_0])^H\big]\end{aligned}$$

Model update:

$$\begin{aligned}\hat{\mathbf{x}}^a_{l,n|n-1} &= \mathbf{A}^a_n \hat{\mathbf{x}}^a_{l,n-1|n-1} \\ \hat{\mathbf{M}}^a_{l,n|n-1} &= \mathbf{A}^a_n \hat{\mathbf{M}}^a_{l,n-1|n-1} \mathbf{A}^{aH}_n + \mathbf{C}_{\boldsymbol{\nu}^a_n}\end{aligned}$$

Measurement update:

$$\begin{aligned}\hat{\mathbf{M}}^{a^{-1}}_{l,n|n} &= \hat{\mathbf{M}}^{a^{-1}}_{l,n|n-1} + \sum_{\forall m \in \mathcal{N}_l} \Big(\mathbf{H}^{aH}_{m,n} \mathbf{C}^{-1}_{\omega^a_{m,n}} \mathbf{H}^a_{m,n}\Big) \\ \mathbf{G}^a_{l,n} &= \hat{\mathbf{M}}^a_{l,n|n} \mathbf{H}^{aH}_{l,n} |\mathcal{N}_l| \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{l,n}} \\ \boldsymbol{\phi}^a_{l,n} &= \hat{\mathbf{x}}^a_{l,n|n-1} + \mathbf{G}^a_{l,n}\big(\mathbf{y}^a_{l,n} - \mathbf{H}^a_{l,n} \hat{\mathbf{x}}^a_{l,n|n-1}\big)\end{aligned}$$

Information sharing:

1. Share $\boldsymbol{\phi}^a_{l,n}$ with neighboring nodes.
2. Share $\mathbf{H}^{aH}_{l,n} \mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{l,n}} \mathbf{H}^a_{l,n}$ with neighboring nodes, only if it has changed compared to the previous time instant.

Estimate fusion :

$$\hat{\mathbf{x}}^a_{l,n|n} = \frac{1}{|\mathcal{N}_l|} \sum_{\forall m \in \mathcal{N}_l} \boldsymbol{\phi}^a_{m,n}$$

Furthermore, substituting

$$\begin{aligned}\boldsymbol{\epsilon}^a_{l,n|n-1} =& \mathbf{x}^a_n - \hat{\mathbf{x}}^a_{l,n|n-1} = \mathbf{A}^a_n \mathbf{x}^a_{n-1} + \boldsymbol{\nu}^a_{l,n} - \mathbf{A}^a_n \hat{\mathbf{x}}^a_{l,n-1|n-1} \\ =& \mathbf{A}^a_n \Big(\underbrace{\mathbf{x}^a_{n-1} - \hat{\mathbf{x}}^a_{l,n-1|n-1}}_{\boldsymbol{\epsilon}^a_{l,n-1|n-1}}\Big) + \boldsymbol{\nu}^a_{l,n}\end{aligned}$$

into the expression in (7.8) gives

$$\boldsymbol{\epsilon}^a_{l,n} = \big(\mathbf{I} - \mathbf{G}^a_{l,n}\mathbf{H}^a_{l,n}\big)\, \mathbf{A}^a_n \boldsymbol{\epsilon}^a_{l,n-1|n-1} + \big(\mathbf{I} - \mathbf{G}^a_{l,n}\mathbf{H}^a_{l,n}\big)\, \boldsymbol{\nu}^a_{l,n} - \mathbf{G}^a_{l,n}\boldsymbol{\omega}^a_{l,n}. \tag{7.9}$$

Now, consider the difference between the true augmented state vector and its estimate obtained at node l, given by

$$\boldsymbol{\epsilon}^a_{l,n|n} = \mathbf{x}^a_{l,n} - \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\boldsymbol{\phi}^a_{m,n} = \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\boldsymbol{\epsilon}^a_{m,n|n} \tag{7.10}$$

where replacing (7.9) into (7.10) gives a recursive expression for the augmented state vector estimation error as

$$\begin{aligned}\boldsymbol{\epsilon}^a_{l,n|n} =& \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\left(\mathbf{I}-\mathbf{G}^a_{m,n}\mathbf{H}^a_{m,n}\right)\mathbf{A}^a_n\boldsymbol{\epsilon}^a_{m,n-1|n-1}\\ &+\frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\left(\mathbf{I}-\mathbf{G}^a_{m,n}\mathbf{H}^a_{m,n}\right)\boldsymbol{\nu}^a_{m,n} - \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\mathbf{G}^a_{m,n}\boldsymbol{\omega}^a_{m,n}.\end{aligned} \tag{7.11}$$

From Algorithm 12, we can now substitute

$$\mathbf{G}^a_{m,n}\mathbf{H}^a_{m,n} = \hat{\mathbf{M}}^a_{m,n|n}\underbrace{\mathbf{H}^{aH}_{m,n}|\mathcal{N}_m|\mathbf{C}^{-1}_{\boldsymbol{\omega}^a_{m,n}}\mathbf{H}^a_{m,n}}_{\mathbf{P}_{m,n}}$$

into (7.11) to yield

$$\begin{aligned}\boldsymbol{\epsilon}^a_{l,n|n} =& \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\left(\mathbf{I}-\hat{\mathbf{M}}^a_{m,n|n}\mathbf{P}_{m,n}\right)\mathbf{A}^a_n\boldsymbol{\epsilon}^a_{m,n-1|n-1}\\ &+\frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\left(\mathbf{I}-\hat{\mathbf{M}}^a_{m,n}\mathbf{P}_{m,n}\right)\boldsymbol{\nu}^a_{m,n} - \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\mathbf{G}^a_{m,n}\boldsymbol{\omega}^a_{m,n}.\end{aligned} \tag{7.12}$$

7.4.1 Mean error behavior

Taking the statistical expectation of the expression in (7.12) allows the mean error behavior of the DAQKF to be formulated as

$$\begin{aligned}E[\boldsymbol{\epsilon}^a_{l,n|n}] =& \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\left(\mathbf{I}-\hat{\mathbf{M}}^a_{m,n|n}\mathbf{P}_{m,n}\right)\mathbf{A}^a_n E[\boldsymbol{\epsilon}^a_{m,n-1|n-1}]\\ &+\frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\left(\mathbf{I}-\hat{\mathbf{M}}^a_{m,n}\mathbf{P}_{m,n}\right)E[\boldsymbol{\nu}^a_{m,n}] - \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\mathbf{G}^a_{m,n}E[\boldsymbol{\omega}^a_{m,n}]\end{aligned}$$

where considering that $\boldsymbol{\nu}^a_{m,n}$ and $\boldsymbol{\omega}^a_{m,n}$ are quaternion-valued zero-mean Gaussian random vectors results in

$$E[\boldsymbol{\epsilon}^a_{l,n|n}] = \frac{1}{|\mathcal{N}_l|}\sum_{\forall m\in\mathcal{N}_l}\left(\mathbf{I}-\hat{\mathbf{M}}^a_{m,n}\mathbf{P}_{m,n}\right)\mathbf{A}^a_n E[\boldsymbol{\epsilon}^a_{m,n-1|n-1}]. \tag{7.13}$$

Therefore, given that $\forall m\in\mathcal{N}:\hat{\mathbf{x}}^a_{m,0|0} = E[\mathbf{x}^a_0]$, the expression in (7.13) indicates that the DAQKF operates in an unbiased fashion.

7.4.2 Local mean square error behavior

Given the error of the augmented state vector estimates in the formulation in (7.12), the augmented covariance matrix of the augmented state vector estimates at node l and time instant n can be expressed as

$$\mathbf{C}_{\boldsymbol{\epsilon}^a_{l,n|n}} = E\left[\boldsymbol{\epsilon}^a_{l,n|n}\boldsymbol{\epsilon}^{aH}_{l,n|n}\right] = \mathbf{S}_{l,n}\mathcal{E}_{n-1}\mathbf{S}^H_{l,n} + \mathbf{R}_{l,n}\mathcal{V}_n\mathbf{R}^H_{l,n} + \mathbf{Q}_{l,n}\mathcal{W}_n\mathbf{Q}^H_{l,n} \tag{7.14}$$

where the expressions

$$\begin{aligned}\mathcal{E}_n =& E\left[\left[\boldsymbol{\epsilon}^{aT}_{1,n|n},\ldots,\boldsymbol{\epsilon}^{aT}_{|\mathcal{N}|,n|n}\right]^T\left[\boldsymbol{\epsilon}^{aT}_{1,n|n},\ldots,\boldsymbol{\epsilon}^{aT}_{|\mathcal{N}|,n|n}\right]^*\right]\\ \mathcal{W}_n =& E\left[\left[\boldsymbol{\omega}^{aT}_{1,n},\ldots,\boldsymbol{\omega}^{aT}_{|\mathcal{N}|,n}\right]^T\left[\boldsymbol{\omega}^{aT}_{1,n},\ldots,\boldsymbol{\omega}^{aT}_{|\mathcal{N}|,n}\right]^*\right]\end{aligned}$$

represent respectively the state estimation error and the observation noise cross-covariances between all nodes in the network, whereas

$$\mathcal{V}_n = E\left[\left[\boldsymbol{\nu}^{aT}_{1,n},\ldots,\boldsymbol{\nu}^{aT}_{|\mathcal{N}|,n}\right]^T\left[\boldsymbol{\nu}^{aT}_{1,n},\ldots,\boldsymbol{\nu}^{aT}_{|\mathcal{N}|,n}\right]^*\right] = \begin{bmatrix}\mathbf{C}_{\boldsymbol{\nu}^a_n} & \cdots & \mathbf{C}_{\boldsymbol{\nu}^a_n}\\ \vdots & \ddots & \vdots\\ \mathbf{C}_{\boldsymbol{\nu}^a_n} & \cdots & \mathbf{C}_{\boldsymbol{\nu}^a_n}\end{bmatrix}$$

while

$$\mathbf{S}_{l,n} = \frac{1}{|\mathcal{N}_l|}\begin{bmatrix}\alpha_{l,1}\mathbf{A}^{aH}_n\left(\mathbf{I}-\hat{\mathbf{M}}^a_{1,n|n}\mathbf{P}_{1,n}\right)^H\\ \alpha_{l,2}\mathbf{A}^{aH}_n\left(\mathbf{I}-\hat{\mathbf{M}}^a_{2,n|n}\mathbf{P}_{2,n}\right)^H\\ \vdots\\ \alpha_{l,|\mathcal{N}|}\mathbf{A}^{aH}_n\left(\mathbf{I}-\hat{\mathbf{M}}^a_{|\mathcal{N}|,n|n}\mathbf{P}_{|\mathcal{N}|,n}\right)^H\end{bmatrix}^H$$

$$\mathbf{R}_{l,n} = \frac{1}{|\mathcal{N}_l|}\begin{bmatrix}\alpha_{l,1}\left(\mathbf{I}-\hat{\mathbf{M}}^a_{1,n|n}\mathbf{P}_{1,n}\right)^H\\ \alpha_{l,2}\left(\mathbf{I}-\hat{\mathbf{M}}^a_{2,n|n}\mathbf{P}_{2,n}\right)^H\\ \vdots\\ \alpha_{l,|\mathcal{N}|}\left(\mathbf{I}-\hat{\mathbf{M}}^a_{|\mathcal{N}|,n|n}\mathbf{P}_{|\mathcal{N}|,n}\right)^H\end{bmatrix}^H$$

$$\mathbf{Q}_{l,n} = \frac{1}{|\mathcal{N}_l|}\left[\alpha_{l,1}\mathbf{G}^a_{1,n},\alpha_{1,2}\mathbf{G}^a_{2,n},\cdots,\alpha_{1,|\mathcal{N}|}\mathbf{G}^a_{|\mathcal{N}|,n}\right]$$

with

$$\alpha_{l,m} = \begin{cases}1, & \text{if } m\in\mathcal{N}_l\\ 0, & \text{otherwise.}\end{cases}$$

Therefore, from the following standard assumptions in steady-state Kalman filtering analysis [130, 135]:

1. the state evolution function and observation functions for all the nodes in the sensor network are time invariant, i.e.

$$\mathbf{A}_n^a \to \mathbf{A}^a \text{ and } \forall l \in \mathcal{N} : \mathbf{H}_{l,n}^a \to \mathbf{H}_l^a \text{ as } n \to \infty$$

2. the state evolution and observation noises are stationary, i.e. the covariance matrices $\mathcal{V}_n$, and $\mathcal{W}_n$ become time invariant, or equivalently

$$\mathcal{V}_n \to \mathcal{V} \text{ and } \mathcal{W}_n \to \mathcal{W} \text{ as } n \to \infty$$

3. the matrix pairs $\forall l \in \mathcal{N} : \{\mathbf{A}^a, \mathbf{H}_{l,n}^a\}$ are detectable and the matrix pair $\{\mathbf{A}^a, \mathbf{C}_{\boldsymbol{\nu}^a}^{\frac{1}{2}}\}$ is stabilizable

it follows that for all nodes in the network, $\hat{\mathbf{M}}_{l,n|n}^a$ becomes time invariant resulting in the matrices $\{\mathbf{S}_{l,n}, \mathbf{R}_{l,n}, \mathbf{Q}_{l,n}\}$ also becoming time invariant and therefore $\mathbf{C}_{\boldsymbol{\epsilon}_{l,n|n}^a}$ converges.

Finally, from Algorithm 12 and the expression in (7.14), notice that the correlation between the observational noise at different nodes in the network does not have an effect on $\mathbf{S}_{l,n}$, $\mathbf{R}_{l,n}$, and $\mathbf{Q}_{l,n}$; in addition, $\mathrm{Tr}(\mathbf{C}_{\boldsymbol{\epsilon}_{l,n|n}^a})$ is linearly dependent on $\mathrm{Tr}(\mathbf{Q}_{l,n}\mathcal{W}_n\mathbf{Q}_{l,n}^H)$. Hence, for a constant value of $\mathrm{Tr}(\mathcal{W}_n)$, $\mathrm{Tr}(\mathbf{C}_{\boldsymbol{\epsilon}_{l,n|n}^a})$ is minimized (*cf.* maximized) when the observational noises at different nodes are uncorrelated (*cf.* fully correlated).

7.4.3 Global mean square error behavior

From the recursive expression of the augmented state vector estimation error given in (7.12), the state vector estimation error augmented cross-covariance between all nodes of the network, $\mathcal{E}_n$, can be formulated in a recursive fashion as

$$\mathcal{E}_n = \mathcal{S}_n\mathcal{E}_{n-1}\mathcal{S}_n^H + \mathcal{R}_n\mathcal{V}_n\mathcal{R}_n^H + \mathcal{Q}_n\mathcal{W}_n\mathcal{Q}_n^H \tag{7.15}$$

where

$$\mathcal{S}_n = \begin{bmatrix} \mathbf{S}_{1,n} \\ \vdots \\ \mathbf{S}_{|\mathcal{N}|,n} \end{bmatrix}, \mathcal{R}_n = \begin{bmatrix} \mathbf{R}_{1,n} \\ \vdots \\ \mathbf{R}_{|\mathcal{N}|,n} \end{bmatrix}, \text{ and } \mathcal{Q}_n = \begin{bmatrix} \mathbf{Q}_{1,n} \\ \vdots \\ \mathbf{Q}_{|\mathcal{N}|,n} \end{bmatrix}.$$

Then, if convergence conditions in Section 7.4.2 are satisfied, i.e. $\mathcal{V}_n \to \mathcal{V}$ and $\mathcal{W}_n \to \mathcal{W}$, as a result of local convergence, matrices $\{\mathcal{S}_n, \mathcal{R}_n, \mathcal{Q}_n\}$ become time invariant, that is

$$\lim_{n\to\infty} \mathcal{S}_n = \mathcal{S}, \ \lim_{n\to\infty} \mathcal{R}_n = \mathcal{R}, \text{ and } \lim_{n\to\infty} \mathcal{Q}_n = \mathcal{Q}$$

and $\mathcal{E}_n$ in (7.15) converges, that is, $\mathcal{E}_n \to \mathcal{E}$ as $n \to \infty$. Therefore, the expression in (7.15) simplifies to a quaternion-valued discrete time Lyanpunov equation given by

$$\mathcal{E} = \mathcal{S}\mathcal{E}\mathcal{S}^H + \mathcal{R}\mathcal{V}\mathcal{R}^H + \mathcal{Q}\mathcal{W}\mathcal{Q}^H. \tag{7.16}$$

Invoking the duality between $\mathbb{R}$ and $\mathbb{H}$ established using the expressions in (2.8) and through decomposing the quaternion-valued matrices in (7.16) into their real-valued components, the closed form solution to the equation in (7.16) can be obtained as

$$\mathrm{Vec}(\mathcal{E}^{\mathbb{HR}}) = \left(\mathcal{I} - \mathcal{S}^{\mathbb{HR}} \otimes \mathcal{S}^{\mathbb{HR}}\right)^{-1} \mathrm{Vec}\left(\mathcal{A}^{\mathbb{HR}}\right) \tag{7.17}$$

where $\mathcal{A} = \mathcal{R}\mathcal{V}\mathcal{R}^H + \mathcal{Q}\mathcal{W}\mathcal{Q}^H$ and $\mathcal{I}$ is an identity matrix with the same number of rows as $\mathcal{S}^{\mathbb{HR}} \otimes \mathcal{S}^{\mathbb{HR}}$, while

$$\mathcal{E}^{\mathbb{HR}} = \begin{bmatrix} \mathcal{E}_r & -\mathcal{E}_i & -\mathcal{E}_j & -\mathcal{E}_k \\ \mathcal{E}_i & \mathcal{E}_r & -\mathcal{E}_k & \mathcal{E}_j \\ \mathcal{E}_j & \mathcal{E}_k & \mathcal{E}_r & -\mathcal{E}_i \\ \mathcal{E}_k & -\mathcal{E}_j & \mathcal{E}_i & \mathcal{E}_r \end{bmatrix}$$

with $\mathcal{S}^{\mathbb{HR}}$ and $\mathcal{A}^{\mathbb{HR}}$ defined analogously.

7.5 Confidence measure

Note that the DAQKF, in the formulation presented in Algorithm 12, assumes that all the nodes in the network are estimating the same augmented state vector sequence; however, in many applications this assumption may not hold true. For example, in a three-phase power distribution network faults can change their characteristics as they propagate throughout the network due to the presents of fault mitigation devices, such as compensators. Therefore, in these applications, it becomes necessary for a node to identify other nodes in its neighborhood that are estimating the same augmented state vector sequence. To this end, based on the results of the performance analysis presented in Section 7.4, we introduce a confidence measure that allows each node to decide weather or not the local augmented state vector estimate from a given neighboring node offers a valid update for its own augmented state vector estimate.

Let $\mathbf{x}^a_{l,n}$ and $\mathbf{x}^a_{m,n}$ denote respectively the augmented state vectors of nodes l and m at time instant n. Considering that $\boldsymbol{\epsilon}^a_{l,n|n-1} = \mathbf{x}^a_{l,n} - \hat{\mathbf{x}}^a_{l,n|n-1}$; from Algorithm 12, $\boldsymbol{\epsilon}^a_{l,n|n-1}$ is a zero-mean quaternion-valued Gaussian random vector with the augmented covariance matrix $\hat{\mathbf{M}}^a_{l,n|n-1}$. Therefore, from (7.11) and considering that $\boldsymbol{\omega}^a_{l,n}$ is a quaternion-valued zero-mean Gaussian random vector, $\boldsymbol{\epsilon}^a_{l,n} = \mathbf{x}^a_{l,n} - \boldsymbol{\phi}^a_{l,n}$ will also be a zero-mean quaternion-valued random vector, the augmented covariance matrix of which is given by

$$\mathbf{C}_{\boldsymbol{\epsilon}^a_{l,n}} = \left(\mathbf{I} - \mathbf{G}^a_{l,n}\mathbf{H}^a_{l,n}\right) \hat{\mathbf{M}}^a_{l,n|n-1} \left(\mathbf{I} - \mathbf{G}^a_{l,n}\mathbf{H}^a_{l,n}\right)^H + \mathbf{G}^a_{l,n}\mathbf{C}_{\boldsymbol{\omega}^a_{l,n}}\mathbf{G}^{aH}_{l,n}. \tag{7.18}$$

A measure of difference between the observation at node m and its predicted value given the local estimate at node l can now be defined as

$$\begin{aligned}
\mathbf{r}^a_{(l,m)} =& \mathbf{H}^a_{m,n}\boldsymbol{\phi}^a_{l,n} - \mathbf{y}^a_{m,n} \\
=& \mathbf{H}^a_{m,n}\left(\boldsymbol{\phi}^a_{l,n} - \mathbf{x}^a_{m,n}\right) - \boldsymbol{\omega}^a_{l,n} \\
=& \mathbf{H}^a_{m,n}\Delta\mathbf{x}^a_{(l,m)_n} - \mathbf{H}^a_{m,n}\boldsymbol{\epsilon}^a_{l,n} - \boldsymbol{\omega}^a_{l,n}
\end{aligned} \tag{7.19}$$

where $\Delta x^a_{(l,m)_n}$ denotes the difference between the augmented state vectors at nodes l and m at time instant n. In addition, from the expression in (7.19), note that $\mathbf{r}^a_{(l,m)}$ is also a quaternion-valued Gaussian random vector with the augmented covariance matrix

$$\mathbf{C}_{\mathbf{r}^a_{(l,m)}} = \mathbf{H}^a_{m,n}\mathbf{C}_{\boldsymbol{\epsilon}^a_{l,n}}\mathbf{H}^{aH}_{m,n} + \mathbf{C}_{\boldsymbol{\omega}^a_{l,n}} \tag{7.20}$$

and mean vector $\mathbf{H}^a_{m,n}\Delta\mathbf{x}^a_{(l,m)_n}$, where $\mathbf{C}_{\boldsymbol{\epsilon}^a_{l,n}}$ is given in (7.18). In cases where nodes l and m are estimating the same augmented state vector sequence; $\Delta\mathbf{x}^a_{(l,m)_n} = 0$ and hence $\mathbf{H}^a_{m,n}\Delta\mathbf{x}^a_{(l,m)_n} = 0$. Therefore, the Mahalanobis distance $d = \mathbf{r}^{aH}_{(l,m)}\mathbf{C}^{-1}_{\mathbf{r}^a_{(l,m)}}\mathbf{r}^a_{(l,m)}$ can be used as a confidence measure to indicate if $\mathbf{r}^a_{(l,m)}$ is an outlier (*cf.* not an outlier) for a zero-mean quaternion-valued distribution with the augmented covariance matrix $\mathbf{C}_{\mathbf{r}^a_{(l,m)}}$ indicating that the local estimate at node l offers an invalid (*cf.* valid) update for the augmented state vector estimates at node m given the measurement $\mathbf{y}^a_{m,n}$.

7.6 Simulations

In this section, first the steady-state MSE performance of the developed DAQKF is examined through a generic distributed Kalman filtering example, where the goal it to validate the performance analysis in Section 7.4. Then, the newly developed DAQKF, is applied for frequency estimation in power grids, where nodes in the power distribution network that are operating under similar conditions cooperate to improve their estimate of the system frequency. In addition, the DAQKF is applied for collaborative target tracking, where each node in the network can only observe the bearings of the target and the DAQKF is exploited to force a consensus on the state of the target that accounts for observations at all the nodes in the network. Note that unless stated otherwise, the network of 20 nodes shown in Figure 7.1 was used for simulations.

In all simulations, akin to current real and complex-valued distributed Kalman filtering approaches [127, 128, 130, 131], the assumption was made that communication between the nodes is immediate (without time delay). In particular, in the case of frequency estimation in power distribution networks, the additional assumption was made that the phase shifts between the observed signal at different locations are negligible, a reasonable assumption as usually sensors are distributed over a small area, for example a micro-grid.

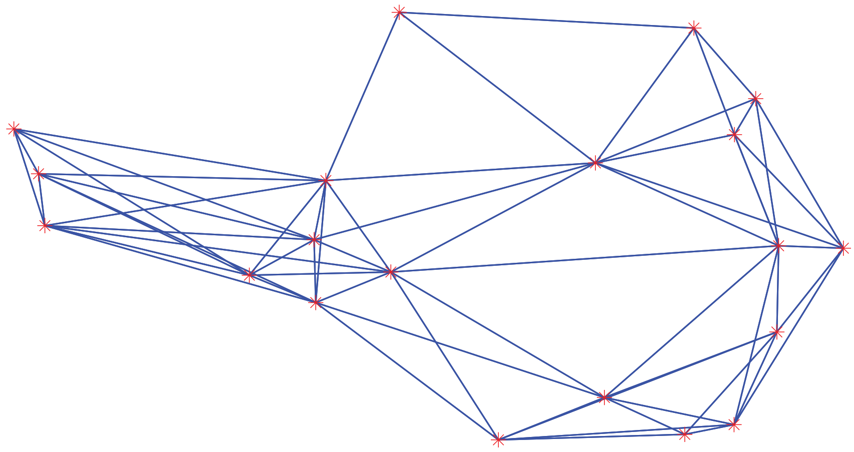

Figure 7.1: The network of 20 nodes used for simulations. Nodes are marked by red "*" and connections are shown with blue lines.

7.6.1 Steady-state performance evaluation

In order to validate the theoretical analysis in Section 7.4, the dynamic system characterized through the following state space equations

$$\forall l \in \mathcal{N}: \begin{cases} \mathbf{x}_n = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}_{n-1} + \begin{bmatrix} 0.5 \\ 1 \end{bmatrix} \nu_n \\ y_{l,n} = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}_n + \omega_{l,n} \end{cases} \tag{7.21}$$

was considered, where ν_n is a unit variance quaternion-valued zero-mean Gaussian second-order proper (circular) noise and $\omega_{l,n}$ are improper quaternion-valued zero-mean Gaussian noise with the following second-order statistics

$$\mathbf{C}_{\omega_{l,n}} = 0.0325,\ \mathbf{R}_{\omega_{l,n}\omega_{l,n}^{i*}} = -0.0075,\ \mathbf{R}_{\omega_{l,n}\omega_{l,n}^{j*}} = -0.0075, \text{ and } \mathbf{R}_{\omega_{l,n}\omega_{l,n}^{k*}} = -0.0075.$$

The developed DAQKF was implemented over the network in Figure 7.1 to track the state vector through observations given in (7.21). The steady-state MSE performance of all nodes in the network obtained through simulations and the theoretical framework in Section 7.4 are shown along side the steady-state MSE performance of the CAQKF in Figure 7.2. Note that the DAQKF achieved an MSE performance close to that of the centralized approach obtained through implementing the CAQKF. In addition, the steady-state MSE obtained through simulations followed those obtained through the theoretical framework in Section 7.4, which verifies the work in that section.

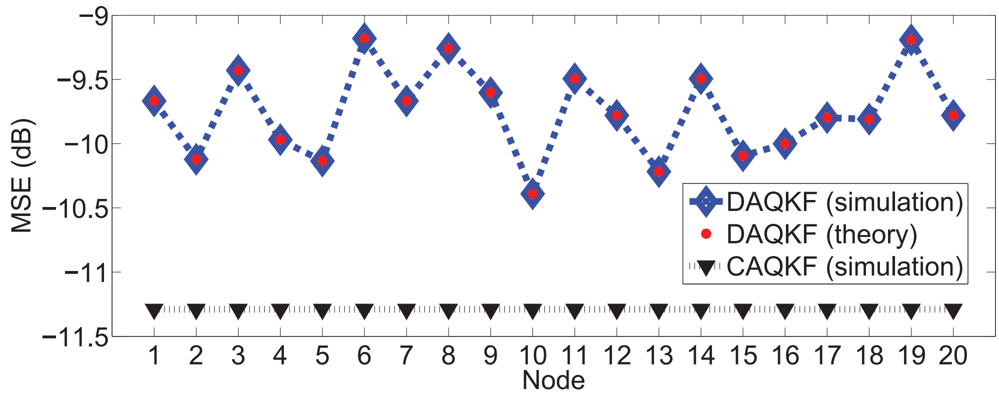

Figure 7.2: The steady-state MSE performance of all nodes in the network implementing the DAQKF and CAQKF.

7.6.2 Smart grid

In the first set of simulations, the three-phase system was considered to be initially operating at its nominal frequency of 50 Hz in a balanced fashion. Then, the system suffered a fault resulting in unbalanced operating conditions characterized by an 80% drop in the amplitude of $v_{a,n}$ and 20 degree shifts in the phases of $v_{b,n}$ and $v_{c,n}$; furthermore, the frequency of the system experienced a step jump of 0.5 Hz. The fault lasted for a short duration and the system returned to its balanced operating condition and its nominal frequency. In Figure 7.3, the estimates of the system frequency obtained through the QFE, that was developed in Chapter 6, implemented using the newly developed DAQKF and the traditional AQKF are shown alongside each other. Note that the estimates of the system frequency obtained through implementing the QFE using the DAQKF have significantly lower steady-state variance as compared to those obtained through implementing the QFE using the AQKF.

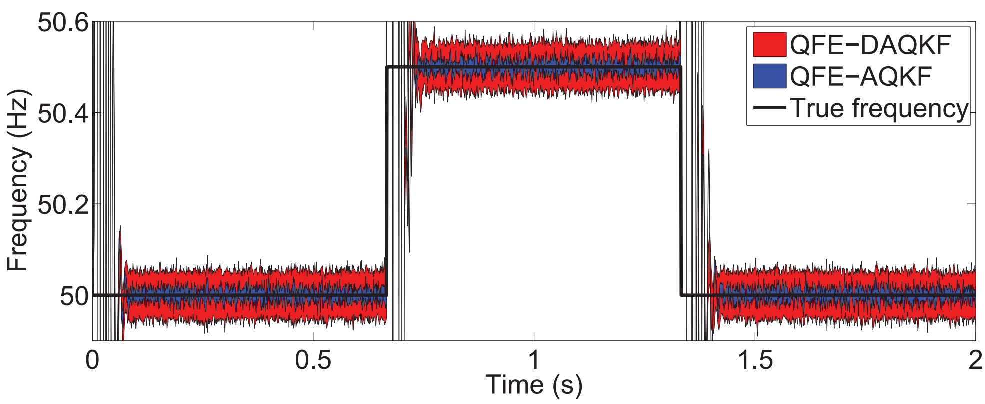

Figure 7.3: Frequency estimation using the QFE implemented through the newly developed DAQKF and the traditional AQKF. The estimate of the system frequency obtained through implementing the QFE by the DAQKF are given in blue and the estimates obtained through implementing the QFE by the AQKF are given in red.

In the second set of simulations, the three-phase system experienced a rising (*cf.* falling) frequency at the rate of 1 Hz/s due to a mismatch between power generation and consumption. In addition, it was assumed that the three-phase system was operating under the same unbalanced conditions characterized in the first simulation. The estimates of the system frequency obtained through the QFE implimented using the DAQKF and AQKF are shown in Figure 7.4. Observe that the QFE implemented through the newly developed DAQKF accurately tracked the system frequency and achieved a lower steady-state variance as compared to the QFE implemented through the AQKF due to cooperation between nodes in the network.

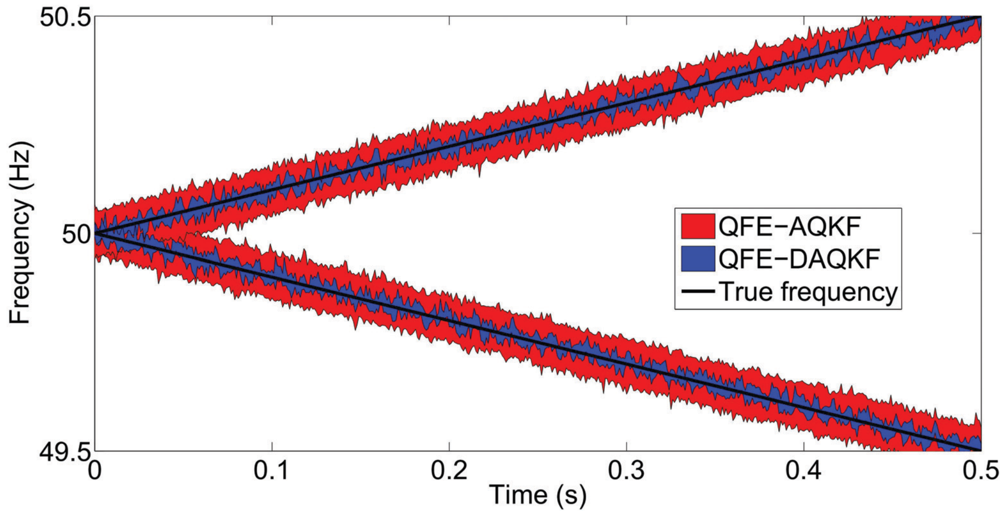

Figure 7.4: Frequency estimation for an unbalanced three-phase system with changing frequency at the rate of 1 Hz/s. The estimate of the system frequency obtained using the QFE implemented through the newly developed DAQKF are in blue and the estimates obtained using the QFE implemented through the AQKF are given in red.

In the third set of simulations, frequency estimation using real-world data recorded from two neighboring nodes in a power distribution network was considered. The recorded data are shown in Figure 7.5, where both nodes suffered a fault 0.1 second after recording started, although Node-2 recovered, Node-1 continued to operate in an unbalanced fashion. The estimates of the system frequency both when the proposed confidence measure was in use and when the proposed confidence measure was ignored are shown in Figure 7.6. Observe that the developed method was able to detect that the nodes are operating under different circumstances and isolated their local estimators preventing bias in the estimated frequency.

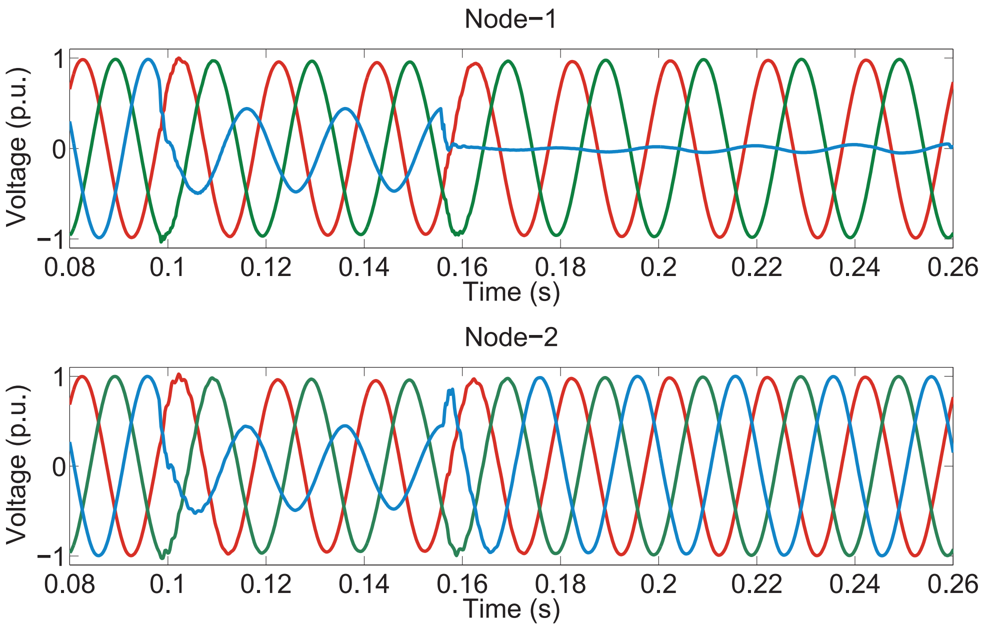

Figure 7.5: Voltage recordings from two neighboring nodes in a real-world power distribution network.

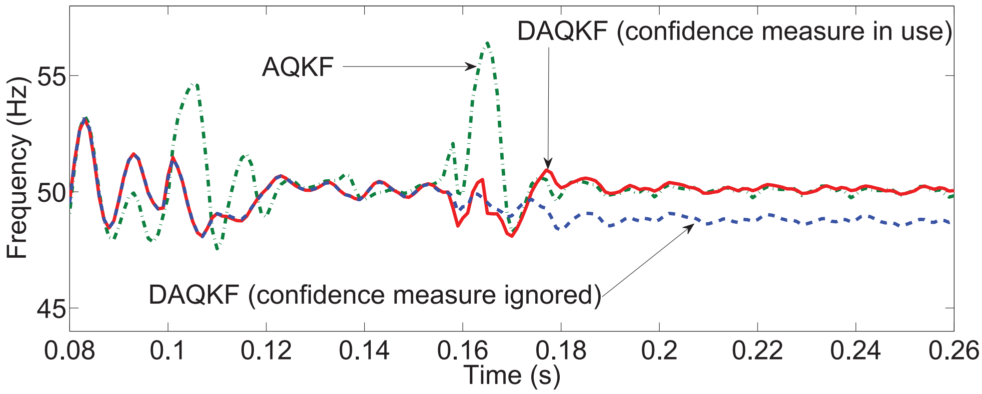

Figure 7.6: Frequency estimation in a power distribution network using real-world data from two neighboring nodes. The estimates obtained using the QFE implemented the AQKF and the DAQKF, both when the proposed confidence measure was in use and when it was ignored, are shown.

The MSE performance of the QFE, implemented through both the AQKF and the newly developed DAQKF, is compared to that of its complex-valued counterparts the CLFE and the CWLFE (see Section 5.3) in Figure 7.7. Notice that the quaternion frequency estimator not only outperformed its linear and widely-linear complex-valued counterparts, but also the unbalanced operating conditions had no significant affect on the performance of the quaternion frequency estimator, a desirable characteristic for frequency estimators in three-phase systems. In addition, employing the developed DAQKF for implementing the QFE resulted in a further reduction of the MSE as compared to the case where the QFE was implemented using the AQKF.

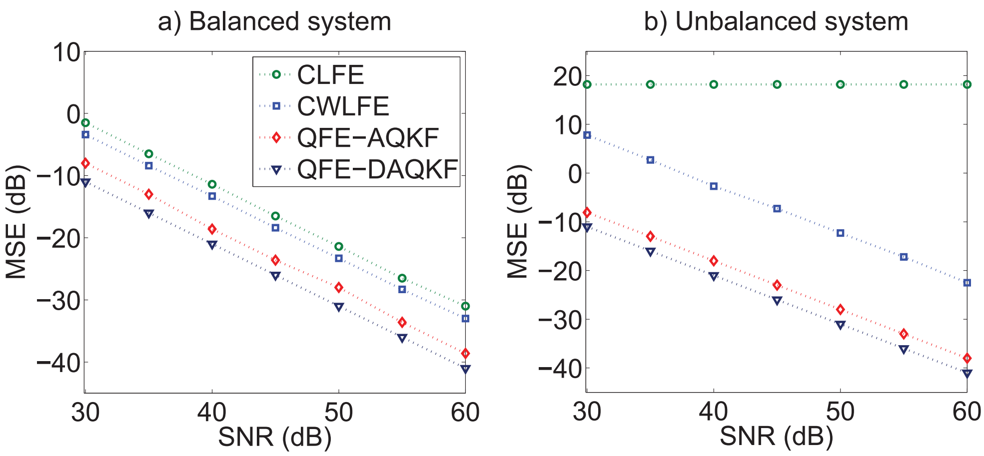

Figure 7.7: MSE performance of the QFE implemented through both the AQKF, QFE-AQKF, and the developed DAQKF, QFE-DAQKF, are compared to its complex-valued counterparts: a) balanced three-phase system, b) unbalanced three-phase system characterized by an 80% drop in the amplitude of $v_{a,n}$ and 20 degree shifts in the phases of $v_{b,n}$ and $v_{c,n}$.

7.6.3 Collaborative target tracking

In this section, we consider the problem of tracking the position of a maneuvering target where the sensors can only measure the bearings of the target. Commonly referred to as bearings-only tracking, this problem is often encountered in passive radar or sonar tracking applications. Since none of the nodes have access to the range of the target, arriving at a unique solution using only the information available to one node is not possible. A solution to this problem is given in [87] using a quaternion Kalman filter that combines the observations of two sensors in order to locate the target through triangulation; however, the results are computationally expensive and are not expandable for implementation over sensor networks.

Taking into account that the developed DAQKF operates akin to a CAQKF that has access to observations from its neighboring nodes, in the solution designed here, the proposed DAQKF is implemented in the sensor network where the diffusion of local estimates is exploited to force the nodes to consent to a unique solution, based on observations of all nodes in the network. The state vector of such a distributed Kalman filter is given by

$$\mathbf{x}_n = \begin{bmatrix} ix_n + jy_n + kz_n \\ i\dot{x}_n + j\dot{y}_n + k\dot{z}_n \end{bmatrix}$$

with $\{x_n, y_n, z_n\}$ and $\{\dot{x}_n, \dot{y}_n, \dot{z}_n\}$ denoting the location and speed of the target on the X, Y, and Z axis. The augmented state evolution function is a block diagonal matrix,

$\mathbf{A}_n^a = \text{block-diag}(\mathbf{A})$, with $\mathbf{A}$ given by

$$\mathbf{A} = \begin{bmatrix} 1 & \Delta T \\ 0 & 1 \end{bmatrix}$$

where ΔT denotes the sampling interval. The state involution noise is in the form

$$\boldsymbol{\nu}_n^a = \mathbf{B}^a \boldsymbol{\eta}_n^a$$

where $\mathbf{B}^a = \text{block-col}(\mathbf{B})$ is a block column matrix with

$$\mathbf{B} = \begin{bmatrix} \frac{1}{2}\left(\Delta T\right)^2 \\ \Delta T \end{bmatrix}$$

while $\boldsymbol{\eta}_n^a$ is a general zero-mean quaternion-valued Gaussian noise. The bearings of the target as measured at sensor l in the network now becomes

$$y_{l,n} = \frac{ix_n + jy_n + kz_n - L_{s_l}}{|ix_n + jy_n + kz_n - L_{s_l}|} + \omega_{l,n}$$

where L_{s_l} denotes the location of sensor l, while $\omega_{l,n}$ represents the observational noise. Finally, it is important to note that since three-dimensional data has been modeled as pure quaternion signals, the state evolution and observation noise sequences are aslo pure quaternions and hence improper quaternion-valued random variables, which further emphasizes the importance of considering the state evolution and observation functions in their augmented formulation in these applications.

The network shown in Figure 7.1 with its nodes distributed uniformly inside a $24 \times 24 \times 24$ cube was used to track a target moving inside the cube through bearings-only measurements. The sampling interval was set to $\Delta T = 0.04$ s, whereas η_n was a zero-mean quaternion-valued Gaussian noise with the following statistics

$$\mathbf{C}_{\nu_n} = 10,\ \mathbf{R}_{\nu_n \nu_n^{i*}} = -3.3,\ \mathbf{R}_{\nu_n \nu_n^{j*}} = -3.3, \text{ and } \mathbf{R}_{\nu_n \nu_n^{k*}} = -3.3.$$

while the observational noise for node all nodes the network was selected as a zero-mean quaternion-valued Gaussian noise with $\mathbf{C}_{\omega_{l,n}} = 0.0001$ and pseudo-covariances given by

$$\mathbf{R}_{\omega_{l,n}\omega_{l,n}^{i*}} = -0.000033,\ \mathbf{R}_{\omega_{l,n}\omega_{l,n}^{j*}} = -0.000033, \text{ and } \mathbf{R}_{\omega_{l,n}\omega_{l,n}^{k*}} = -0.000033.$$

The estimate of the location of the target at a node in the network is shown in Figure 7.8. Note that the proposed algorithm accurately tracks the location of the target.

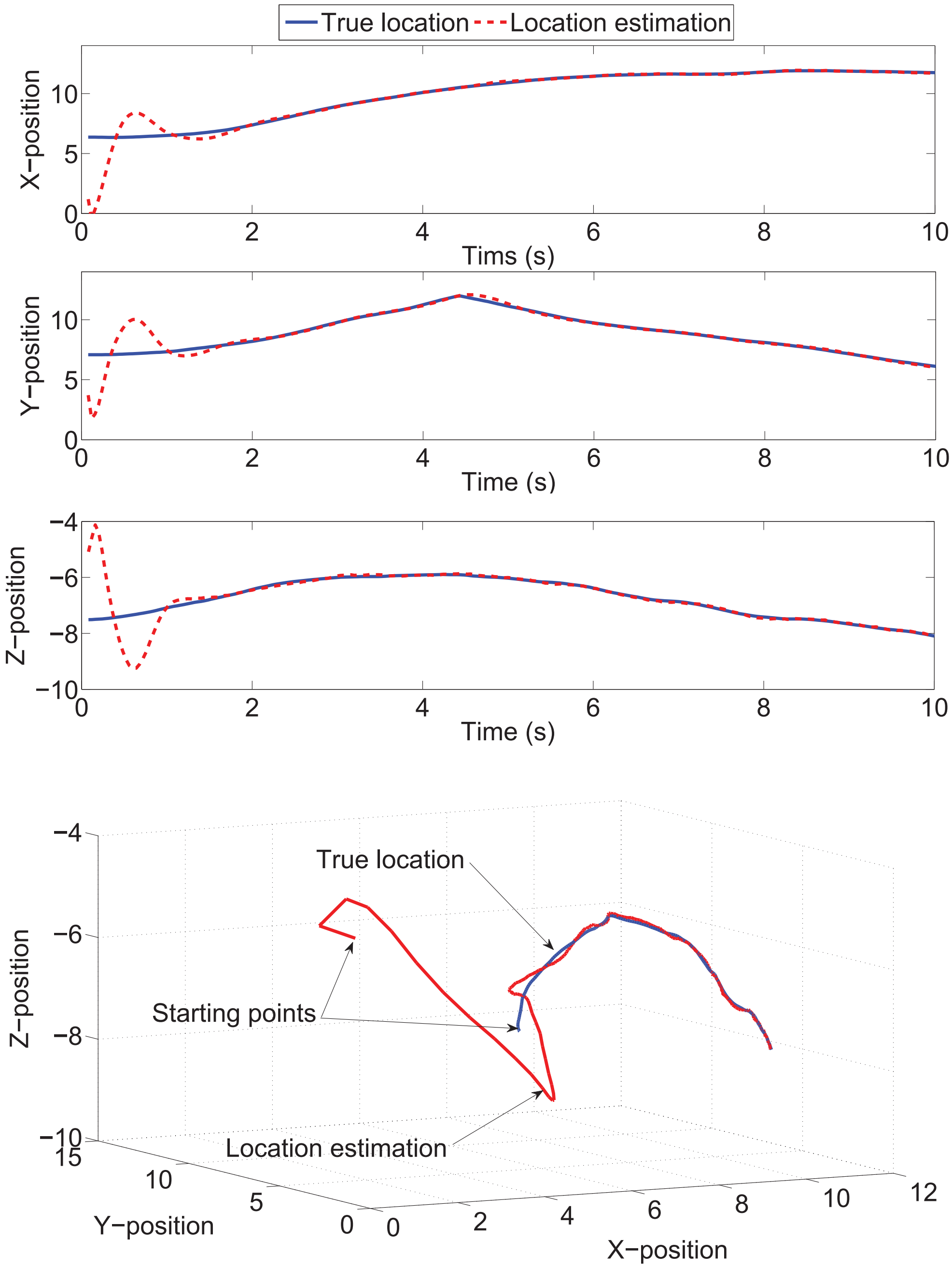

Figure 7.8: Collaborative target tracking using bearings-only measurements. Position of the target and its estimates on the X, Y, and Z axis are shown in the top three graphs, while the bottom graph shows the location of the target and its estimate in the three-dimensional space.

7.7 Conclusion

A distributed quaternion Kalman filter has been developed for distributed sequential state estimation in sensor networks. This has been achieved through decomposing the operations of the centralized quaternion Kalman filter in such a fashion that they can be performed by individual nodes in the network so that the final state vector estimate can be obtained by averaging local estimates calculated at each node. The proposed algorithm differs from existing diffusion Kalman filtering techniques in that it does not require mixing coefficients for averaging local estimates and that the proposed algorithm takes into account the effect of diffusing local estimates on the *a posteriori* estimate of the augmented covariance matrix of the state vector estimation error. The performance of the developed algorithm has been analyzed and a recursive expression for the estimation error has been derived. Finally, the developed algorithm has been used for estimating the fundamental frequency of three-phase power distribution networks and for collaborative target tracking where each sensor can only observe the bearings of the target.

Chapter 8

The Augmented Quaternion Characteristic Function and Filtering of Quaternion Stable Random Signals

8.1 Overview

In recent years, quaternion-valued signal processing has proven to be advantageous in a number of engineering applications. However, these applications have also revealed that in many scenarios, in order to accurately model the underlying physical signal, it is important to include heavy tailed non-Gaussian elements. Classified using their characteristic functions, α-stable random processes, that indeed include the Gaussian case, have proven to be a useful tool in statistical modeling of heavy tailed random signals due to their stable property and the generalized central limit theorem. In this chapter, we first investigate the properties of the characteristic functions of quaternion-valued random variables, such as the link between the characteristic function and circularity; then, the adaptive filtering of elliptically-contoured quaternion-valued stable random processes is considered where a quaternion-valued particle filtering algorithm is proposed.

8.2 Introduction

In most applications assuming a Gaussian model for the signal is justifiable as it often leads to computationally efficient and analytically tractable signal processing algorithms. However, there are many applications ranging from finance to engineering where the observed data and/or noise exhibits sharp spikes and therefore their resulting distribution does not decay as fast as the Gaussian case [136–139]. In these applications, outliers cannot be dealt with as mistakes, since they constitute an integral part of the signal. For modeling these types of signals, random processes with stable distributions that admit a generalized version of the central limit theorem have proven to be a useful tool [138, 139].

A real-valued random vector, $\boldsymbol{X}$, is defined as α-*stable* if and only if for any $m > 2$ there exists $\alpha \in (0, 2]$ such that

$$\boldsymbol{X}^{(1)} + \cdots + \boldsymbol{X}^{(m)} = m^{1/\alpha}\boldsymbol{X} \tag{8.1}$$

where $\{\boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(m)}\}$ are independent copes of $\boldsymbol{X}$ and their summation is equivalent in distribution to that of $m^{1/\alpha}\boldsymbol{X}$ [138]. In the real-valued univariate case, α-stable random variables admit characteristic functions in the form of [138]

$$\boldsymbol{\Phi}_X(s) = E\left[e^{\zeta sx}\right] = e^{\zeta as - \gamma|s|^\alpha(1-\eta\beta \mathrm{sign}(s) f(s,\alpha))}$$

where

$$f(s,\alpha) = \begin{cases} \tan(\alpha\pi/2) & \alpha \neq 1 \\ \frac{2}{\pi}\log(|s|) & \alpha = 1 \end{cases} \qquad \text{and} \qquad \mathrm{sign}(s) = \begin{cases} 1 & s > 0 \\ 0 & s = 0 \\ -1 & s < 0 \end{cases}$$

whereas $\zeta^2 = -1$, while $\gamma > 0$, $0 < \alpha \leq 2$, and $-1 < \beta < 1$. With regards to real-valued univariate random variables the following remarks can be made [137–139]:

1. As the parameter γ increases the distribution of X becomes increasingly dispersed. This is akin to the effect of variance on Gaussian random variables; indeed, when $\alpha = 2$ the random variable X is Gaussian and its variance is given by 2γ.

2. The parameter β controls the skewness of the distribution of X, where $\beta = 0$ provides for a symmetric distribution about the center a, whereas $\beta > 0$ (*cf.* $\beta < 0$) result in a random variable that is skewed to the right (*cf.* left) with the direction of skewness reversed for the special case of $\alpha = 1$.

3. The shift parameter, a, shifts the distribution of the random variable to the right (*cf.* left) if $a > 0$ (*cf.* $a < 0$) akin to the effect of mean in Gaussian random variables; indeed, for the case that $\alpha > 1$ and $\beta = 0$, then, a is the mean of the random variable.

4. The parameter, α, is referred to as the characteristic exponent and controls the heaviness of the tails of the density function. A small positive value of α indicates severe impulsiveness, and thus tails are heavier, while a value of α close to 2 indicates more Gaussian type behavior. A value of $\alpha = 1$ (*cf.* $\alpha = 2$) corresponds to the Cauchy (*cf.* Gaussian) distribution.

The so-called stable property of α-stable random variables and the generalized central limit theorem, which states that if a limit exists for a sum of independent identically distributed (i.i.d) random variables, then this limit must be a stable distribution, makes α-stable random variable the ideal candidate for modeling heavy tailed random processes, where mostly symmetric α-stable (SαS) random variables have become popular [138, 140, 141] due to the fact that their characteristic function simplifies into

$$\mathbf{\Phi}_X(s) = E\left[e^{\zeta s x}\right] = e^{-\gamma|s|^{\alpha}}. \tag{8.2}$$

Notice that the characteristic function in (8.2) only has well defined derivatives of all orders at $s = 0$ when $\alpha = 2$; therefore, from the class of stable random variables only the Gaussian case has mean, variance, and higher-order statistical moments, whereas in general, α-stable random variables, excluding the Gaussian case, only have moments of order less that α [137]. Thus, signal processing techniques based on minimizing the second-oder moment of an error measure, such as the LMS and NLMS algorithms, do not fair well when applied to the generality of α-stable random variables. To this end, extensions based on adaptively minimizing the p^{th} order moments of a given error measure, with $p \leq \alpha$, have become popular due to their computational efficiency [142–144]. However, the main drawback of these techniques is lack of a convergence bound on their adaptation gain. Although rigorous convergence analysis for steepest-decent algorithms employing first, second, and forth order of the absolute error measure exist [145]; these results were obtained under the assumption that the signals of interest have finite variances and are not expandable to the case of stable random signals.

In the real-valued univariate case, stable distributions are now mostly accessible as there are reliable techniques to compute their densities, distribution functions, and estimate their parameters based on empirical characteristic functions and fractional moments [138, 146]. However, due to the complicated dependence structure of their components, the body of work regarding stable complex and quaternion-valued random signals is limited. In this chapter, characteristic functions of quaternion-valued random variables are considered from a geometric point of view, their link to statistical moments, and the design of elliptical quaternion-valued random variables. Then, adaptive filtering of elliptically distributed quaternion-valued stable random signals is investigated, where a quaternion-valued particle filtering algorithm is proposed.

8.3 Characteristic functions of quaternion-valued random variables

With the exception of some special cases, such as the Cauchy and Gaussian distributions, in general, a closed form expression for the pdf of stable random variables dose not exist, resulting in stable random variables being classified using their characteristic functions. Therefore, statistical analysis of quaternion-valued random variables is considered through the application of their characteristic functions. In this section, characteristic functions of quaternion-valued random variables are considered and their link to statistical moments of quaternion-valued random variables is investigated. Moreover, the results are exploited for the design of circular quaternion-valued random variables.

8.3.1 The augmented quaternion characteristic function

Consider the quaternion-valued random vector $\boldsymbol{Q}$ where the joint statistical information of the real-valued components of $\boldsymbol{Q}$ can fully be describe through their joint characteristic function given by

$$\boldsymbol{\Phi}_{\boldsymbol{Q}}(\mathbf{s}_r, \mathbf{s}_i, \mathbf{s}_j, \mathbf{s}_k) = E\left[e^{\zeta(\mathbf{s}_r^T\mathbf{q}_r + \mathbf{s}_i^T\mathbf{q}_i + \mathbf{s}_j^T\mathbf{q}_j + \mathbf{s}_k^T\mathbf{q}_k)}\right] \tag{8.3}$$

with $\zeta \in \mathbb{H}$ such that $\zeta^2 = -1$. Now, following the same approach as the $\mathbb{HR}$-calculus and through exploiting the transformation in (2.8) the characteristic function in (8.3) can be expressed directly in the quaternion domain as

$$\boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a) = E\left[e^{(\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{q}^a)}\right] \tag{8.4}$$

where $\mathbf{s} = \mathbf{s}_r + i\mathbf{s}_i + j\mathbf{s}_j + k\mathbf{s}_k$. Note that in (8.4) the quaternion random vector must be used in its augmented form and therefore the expression in (8.4) is referred to as the augmented quaternion characteristic function (AQCF).

8.3.2 Geometric interpretation of the AQCF

Considering the expressions in (2.8) it can be shown that $\mathbf{s}^{aH}\mathbf{q}^a$ has vanishing imaginary components[1]; therefore, the transformation $e^{(\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{q}^a)}$, for a given value of $\mathbf{s} \in \mathbb{H}^N$, maps the quaternion random vector onto the perimeter of a unit circle in the 1-ζ plane of $\mathbb{H}^N$ that is centered around the origin and hereafter is referred to as "*the unit circle*". Thus, if the quaternion random vector, $\boldsymbol{Q}$, has a positive mass at $\mathbf{q}$, in the distribution of the random variable $e^{(\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{q}^a)}$, this mass will appear at an angle of $\frac{1}{4}\mathbf{s}^{aH}\mathbf{q}^a$ from the real axis on the perimeter of the unit circle in the 1-ζ plane. This is equivalent to taking the distribution of the real-valued random variable $\frac{1}{4}\mathbf{s}^{aH}\mathbf{q}^a$ and warping it around the unit circle in the 1-ζ plane; hence, $E[e^{(\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{q}^a)}]$ will represent the center of mass of such a "*warped-around*" distribution. Regarding the geometric interpretation of the AQCF the following five notable remarks can be made:

[1]Note that $\mathbf{s}^{aH}\mathbf{q}^a = (\mathbf{s}^H\mathbf{q}) + (\mathbf{s}^H\mathbf{q})^i + (\mathbf{s}^H\mathbf{q})^j + (\mathbf{s}^H\mathbf{q})^k = 4\Re(\mathbf{s}^H\mathbf{q}) \in \mathbb{R}$

1. Note that as a result of the transform $e^{(\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{q}^a)}$ being bounded, lying on or within the unit circle, the AQCF is guaranteed to exist for all values of $\mathbf{s} \in \mathbb{H}^N$ regardless of the distribution of $\boldsymbol{Q}$.

2. Note that if the quaternion random vector has a symmetric distribution[2]; then, its corresponding AQCF will have a vanishing imaginary component as the center of mass of its corresponding warped-around distribution will only lie on the real line. Following the same logic, if the distribution of a given quaternion random variable is symmetric along a line, its corresponding AQCF will have vanishing imaginary component along the same line.

3. In the degenerate case where all the weight of $\boldsymbol{Q}$ is concentrated at one point, the center of mass for $e^{(\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{q}^a)}$ lies exclusively on the perimeter of the unit circle. In this case, the AQCF is given by $\boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a) = \cos(\frac{1}{4}\mathbf{s}^{aH}\mathbf{q}^a) + \zeta\sin(\frac{1}{4}\mathbf{s}^{aH}\mathbf{q}^a)$ that represents a periodic function, where the principal direction of periodicity is parallel to the mean vector; however, it is important to note that a random quaternion variable with a purely periodic AQCF is not necessarily degenerate.

4. Considering that a quaternion random vector $\boldsymbol{Q}_{\boldsymbol{\varrho}}$ with mean vector $\boldsymbol{\varrho}$ can be decomposed into $\boldsymbol{Q}_{\boldsymbol{\varrho}} = \boldsymbol{Q} + \boldsymbol{\varrho}$ where $\boldsymbol{Q}$ is a zero-mean random vector; therefore, the AQCF of $\boldsymbol{Q}_{\boldsymbol{\varrho}}$ is given by $\boldsymbol{\Phi}_{\boldsymbol{Q}^a_{\boldsymbol{\varrho}}}(\mathbf{s}^a) = e^{(\frac{\zeta}{4}\mathbf{s}^{aH}\boldsymbol{\varrho}^a)}\boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)$ which in essence is the AQCF of $\boldsymbol{Q}$ modulated by $e^{(\frac{\zeta}{4}\mathbf{s}^{aH}\boldsymbol{\varrho}^a)}$, the AQCF of the degenerate quaternion random vector with all its mass concentrated at $\boldsymbol{\varrho}$.

5. Considering the inherent duality between the AQCF and the joint characteristic function of its real-valued components, established in (8.3) and (8.4), it can be traced that $|\boldsymbol{\Phi}_{\boldsymbol{Q}^a}| \to 0$ as $\|\mathbf{s}^a\|_2 \to \infty$.

6. At $\mathbf{s}^a = 0$ we have $\mathbf{s}^{aH}\mathbf{q}^a = 0$ resulting in $\boldsymbol{\Phi}_{\boldsymbol{Q}^a}(0) = E[e^{(0)}] = 1$ regardless of the distribution of $\boldsymbol{Q}$.

In order to clarify the above mentioned remarks, the real and imaginary components of the AQCF of the degenerate distribution $P_{\boldsymbol{Q}^a}(q = \frac{4}{\sqrt{3}}(1+i+k)) = 1$ is shown in Figure 8.1. Note that both the real and imaginary components are periodic with their principal direction of periodicity along a vector passing through the origin and $1+i+k$; hence, there are no changes in the AQCF with respect to changes on j-axis. The AQCF of a non-circular quaternion Gaussian random variable is shown in Figure 8.2 and the AQCF of the same non-circular quaternion Gaussian random variable with $\varrho = 1+i+k$ added as mean is shown in Figure 8.3, where the modulations introduced as a result of the mean can clearly be observed; in addition, the AQCF of the zero-mean quaternion Gaussian random variable has a vanishing imaginary component as its distribution is symmetric whereas the introduction of a mean results in an imaginary component that

[2]A random quaternion variable is referred to as symmetric if its distribution is symmetric along any straight line passing through the origin.

has the same amplitude envelop as the real component, but is modulated by a wave that is $\pi/2$ out of phase with that modulating the real component.

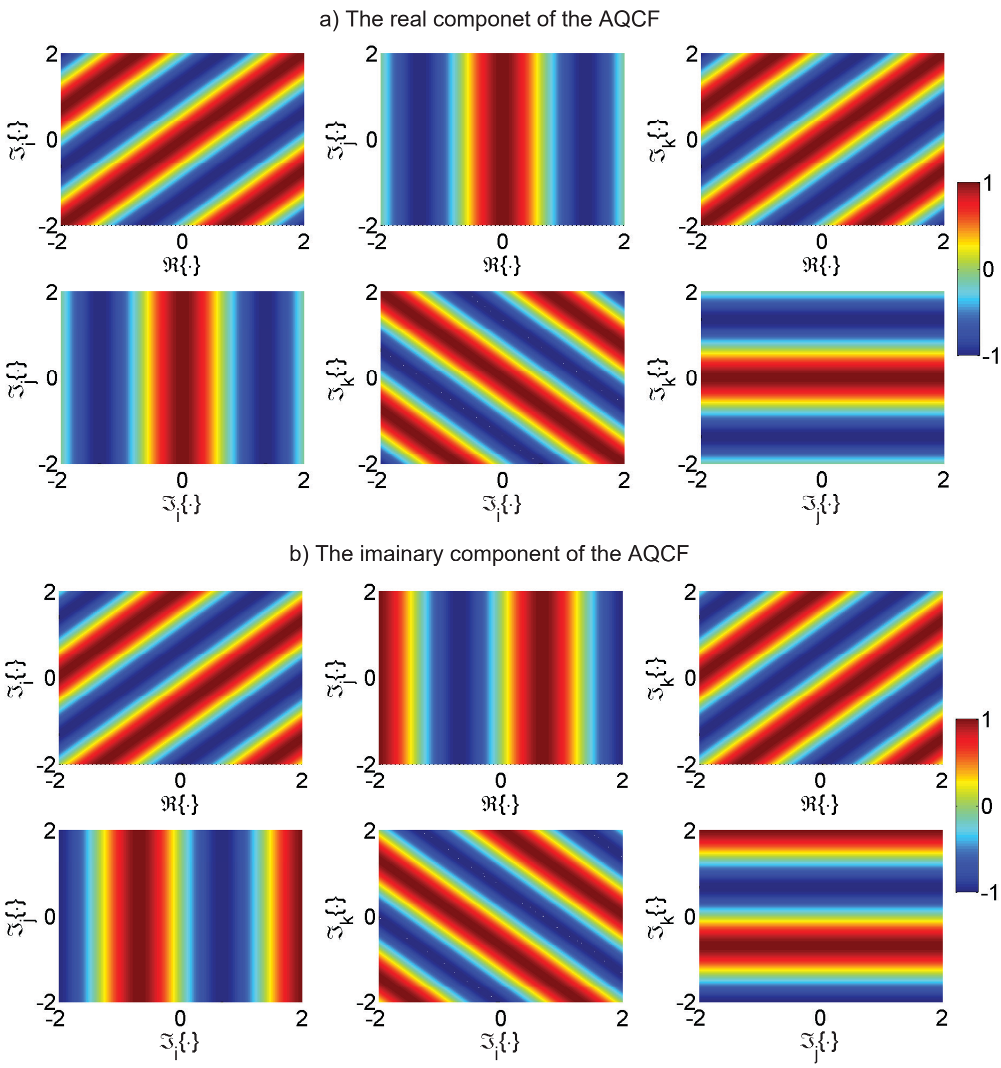

Figure 8.1: Heat map of the AQCF of the degenerate distribution $P_{\boldsymbol{Q}^a}(q = \frac{4}{\sqrt{3}}(1 + i + k)) = 1$ showing variations of the AQCF in six orthogonal planes of $\mathbb{H}$: a) the real component of the AQCF, b) the imaginary component of the AQCF.

8.3.3 The relation between the AQCF and statistical moments

Consider the AQCF in the formulation given in (8.4), where replacing the exponential with its power series representation yields

$$\begin{aligned}\boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a) = E\left[e^{\left(\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{q}^a\right)}\right] =& E\left[\sum_{m=0}^{\infty}\frac{\zeta^m\left(\mathbf{s}^{aH}\mathbf{q}^a\right)^m}{(4^m)m!}\right] \\ =& 1 + E\left[\zeta\frac{\mathbf{s}^{aH}\mathbf{q}^a}{4}\right] - E\left[\frac{\left(\mathbf{s}^{aH}\mathbf{q}^a\right)^2}{32}\right] - E\left[\zeta\frac{\left(\mathbf{s}^{aH}\mathbf{q}^a\right)^3}{384}\right]\cdots\end{aligned} \tag{8.5}$$

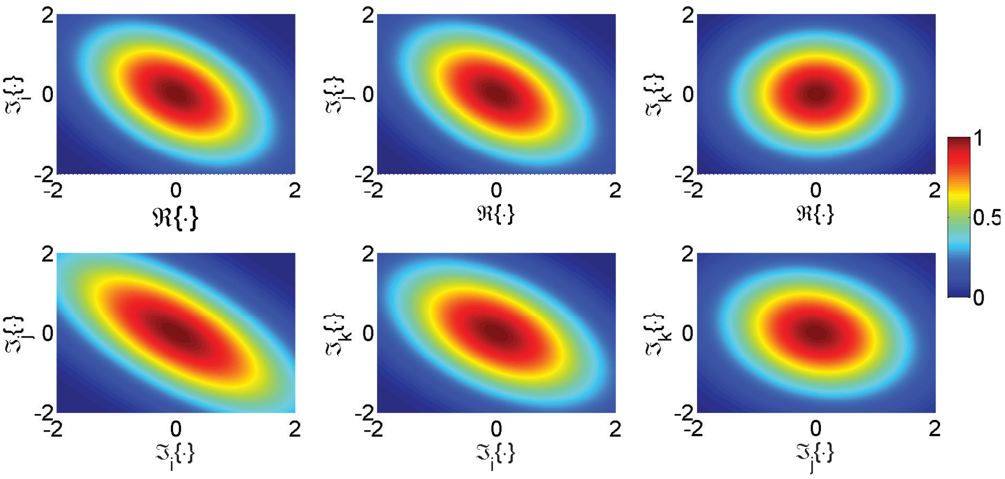

Figure 8.2: Heat map of the AQCF of a zero-mean non-circular quaternion-valued Gaussian distribution showing the real component of the variations of the AQCF in six orthogonal planes of $\mathbb{H}$. Note that the AQCF of this random variable has a vanishing imaginary component.

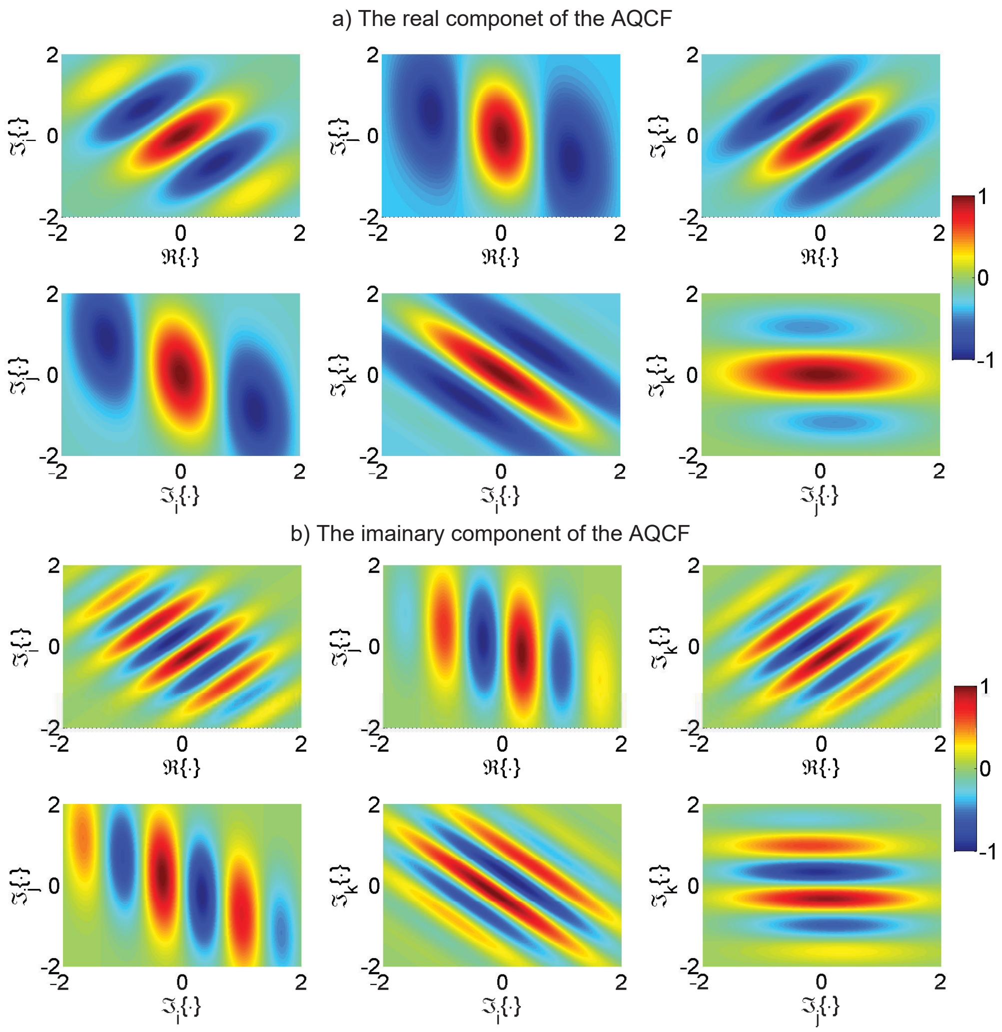

Figure 8.3: Heat map of the AQCF of a non-zero-mean non-circular quaternion-valued Gaussian distribution showing variations of the AQCF in six orthogonal planes of $\mathbb{H}$: a) the real component of the of the AQCF, b) the imaginary component of the AQCF.

The $\mathbb{HR}$-calculus can now be employed to calculate the first-order derivatives of $\boldsymbol{\Phi}_{\boldsymbol{Q}^a}$, where from the expression in (8.5) we have

$$\begin{aligned}\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{*\zeta_1}} =& \frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{aH}\mathbf{q}^a}\frac{\partial \mathbf{s}^{aH}\mathbf{q}^a}{\partial \mathbf{s}^{*\zeta_1}} = E\left[\sum_{m=0}^{\infty}\frac{\zeta^{m+1}\left(\mathbf{s}^{aH}\mathbf{q}^a\right)^m}{(4^{m+1})m!}\mathbf{q}^{\zeta_1}\right] \\ =& E\left[\zeta\frac{\mathbf{q}^{\zeta_1}}{4}\right] - E\left[\frac{\mathbf{s}^{aH}\mathbf{q}^a}{16}\mathbf{q}^{\zeta_1}\right] - E\left[\zeta\frac{\left(\mathbf{s}^{aH}\mathbf{q}^a\right)^2}{128}\mathbf{q}^{\zeta_1}\right]\cdots\end{aligned} \tag{8.6}$$

with $\zeta_1 \in \{1, i, j, k\}$ and when evaluated at $\mathbf{s}^a = 0$ gives

$$\left.\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{*\zeta_1}}\right|_{\mathbf{s}^a=0} = E\left[\zeta\frac{\mathbf{q}^{\zeta_1}}{4}\right]$$

which is the involution of the mean of the random variable $\boldsymbol{Q}$ around ζ_1 multiplied by $\zeta/4$; therefore, we have

$$E\left[\mathbf{q}\right] = \left(\frac{4}{\zeta}\left.\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{*\zeta_1}}\right|_{\mathbf{s}^a=0}\right)^{\zeta_1}.$$

In addition, in a similar manner, it can be shown that

$$\left.\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{\zeta_1}}\right|_{\mathbf{s}^a=0} = \frac{1}{2}E\left[\zeta\frac{\mathbf{q}^{*\zeta_1}}{4}\right] \quad \text{and} \quad E\left[\mathbf{q}\right] = \left(\frac{8}{\zeta}\left.\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{\zeta_1}}\right|_{\mathbf{s}^a=0}\right)^{*\zeta_1}.$$

The second-order derivatives of $\boldsymbol{\Phi}_{\boldsymbol{Q}^a}$ can be expressed as

$$\frac{\partial^2 \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{*\zeta_1}\partial \mathbf{s}^{*\zeta_2}} = \frac{\partial}{\partial \mathbf{s}^{*\zeta_1}}\left(\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{*\zeta_2}}\right) = \frac{\partial}{\partial \mathbf{s}^{aH}\mathbf{q}^a}\left(\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{aH}\mathbf{q}^a}\frac{\partial \mathbf{s}^{aH}\mathbf{q}^a}{\partial \mathbf{s}^{*\zeta_2}}\right)\frac{\partial \mathbf{s}^{aH}\mathbf{q}^a}{\partial \mathbf{s}^{*\zeta_1}} \tag{8.7}$$

where $\zeta_1, \zeta_2 \in \{1, i, j, k\}$ and upon replacing (8.6) into (8.7) gives

$$\frac{\partial^2 \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{*\zeta_1}\partial \mathbf{s}^{*\zeta_2}} = E\left[\sum_{m=0}^{\infty}\frac{-\zeta^m\left(\mathbf{s}^{aH}\mathbf{q}^a\right)^m}{(4^{m+2})m!}\mathbf{q}^{\zeta_2}\mathbf{q}^{\zeta_1^T}\right]$$

that when evaluated at $\mathbf{s}^a = 0$ yields

$$\left.\frac{\partial^2 \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^{*\zeta_1}\partial \mathbf{s}^{*\zeta_2}}\right|_{\mathbf{s}^a=0} = \frac{-1}{16}E\left[\mathbf{q}^{\zeta_2}\mathbf{q}^{\zeta_1 T}\right]$$

which is the cross-correlation between $\mathbf{q}^{\zeta_1}$ and $\mathbf{q}^{\zeta_2}$ multiplied by $(\zeta/4)^2$. Therefore, following the same procedure, the augmented covariance matrix of $\boldsymbol{Q}$ can now be expressed as

$$\left.\frac{\partial^2 \boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)}{\partial \mathbf{s}^a\partial \mathbf{s}^{*a}}\right|_{\mathbf{s}^a=0} = \frac{1}{32}E\left[\mathbf{q}^a\mathbf{q}^{aH}\right] = \frac{1}{32}\mathbf{C}_{\mathbf{q}^a}. \tag{8.8}$$

Note that all higher order statistics of $\boldsymbol{Q}$ can be found analogously.

8.3.4 The relation between the AQCF and circularity

Consider a quaternion-valued random vector, $\boldsymbol{Q}$, with i.i.d components, where from (8.4) and (8.3) it is straightforward to show

$$\boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a) = \boldsymbol{\Phi}_{\boldsymbol{Q}_r}(\mathbf{s}_r)\boldsymbol{\Phi}_{\boldsymbol{Q}_i}(\mathbf{s}_i)\boldsymbol{\Phi}_{\boldsymbol{Q}_j}(\mathbf{s}_j)\boldsymbol{\Phi}_{\boldsymbol{Q}_k}(\mathbf{s}_k)$$

moreover, making the assumption that $\boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a) = g(d)$, where $g(\cdot) : \mathbb{R} \rightarrow \mathbb{C}$ and

$$d = \sqrt{\mathbf{s}_r^T \mathbf{G}_r \mathbf{s}_r + \mathbf{s}_i^T \mathbf{G}_i \mathbf{s}_i + \mathbf{s}_j^T \mathbf{G}_j \mathbf{s}_j + \mathbf{s}_k^T \mathbf{G}_k \mathbf{s}_k} \tag{8.9}$$

results in

$$\frac{\partial g(d)}{\partial \mathbf{s}_r} = \frac{\partial g(d)}{\partial d}\frac{\partial d}{\partial \mathbf{s}_r} = \frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}_r}(\mathbf{s}_r)}{\partial \mathbf{s}_r}\boldsymbol{\Phi}_{\boldsymbol{Q}_i}(\mathbf{s}_i)\boldsymbol{\Phi}_{\boldsymbol{Q}_j}(\mathbf{s}_j)\boldsymbol{\Phi}_{\boldsymbol{Q}_k}(\mathbf{s}_k). \tag{8.10}$$

Through straightforward mathematical manipulations the expression in (8.10) gives

$$\frac{\partial g(d)}{\partial d}\frac{\mathbf{G}_r \mathbf{s}_r}{d} = \frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}_r}(\mathbf{s}_r)}{\partial \mathbf{s}_r}\boldsymbol{\Phi}_{\boldsymbol{Q}_i}(\mathbf{s}_i)\boldsymbol{\Phi}_{\boldsymbol{Q}_j}(\mathbf{s}_j)\boldsymbol{\Phi}_{\boldsymbol{Q}_k}(\mathbf{s}_k) \tag{8.11}$$

where making the assumption that $\mathbf{G}_r$ is positive definite and dividing both sides of the expression in (8.11) by $g(d)$ yields

$$\frac{\partial g(d)}{\partial d}\frac{1}{dg(d)} = \frac{\mathbf{s}_r^T}{\mathbf{s}_r^T \mathbf{G}_r \mathbf{s}_r}\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}_r}(\mathbf{s}_r)}{\partial \mathbf{s}_r}. \tag{8.12}$$

Note that the right hand side of the equation in (8.12) is only dependent on $\mathbf{s}_r$ whereas its left hand side is dependent on d, a function of $\{\mathbf{s}_r,\, \mathbf{s}_i,\, \mathbf{s}_j,\, \mathbf{s}_k\}$; therefore, the right hand side of the equation in (8.12) can only be a constant resulting in

$$\frac{\partial g(d)}{\partial d}\frac{1}{g(d)} = \frac{\partial \ln\left(g(d)\right)}{\partial d} = \underbrace{\left(\frac{\mathbf{s}_r^T}{\mathbf{s}_r^T \mathbf{G}_r \mathbf{s}_r}\frac{\partial \boldsymbol{\Phi}_{\boldsymbol{Q}_r}(\mathbf{s}_r)}{\partial \mathbf{s}_r}\right)}_{\kappa} d \tag{8.13}$$

where solving the differential equation in (8.13) gives[3]

$$g(d) = e^{\left(\frac{\kappa}{2}d^2\right)}. \tag{8.14}$$

Note that the AQCF given in (8.14) is of a form consistent only with the AQCF of quaternion-valued Gaussian variables. Indeed in Appendix D, it is shown that assuming the pdf of $\boldsymbol{Q}$ exists and has i.i.d real-valued components the resulting distribution will be circular if and only if the real-valued components of $\boldsymbol{Q}$ are Gaussian random variables with the same diagonal covariance matrices. Furthermore, notice that the derivatives of the AQCF with respect to the real component were used here; however, the same results

[3] Note that the boundary conditions of the differential equation in (8.13) are set by the fact that for a general quaternion-valued random vector, $\boldsymbol{Q}$, we have $\boldsymbol{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a)|_{\mathbf{s}^a=0} = 1$.

can be obtained by employing the derivatives of the AQCF with respect to any of the real-valued components of the quaternion-valued random vector.

8.3.5 The design of quaternion circular random variables

In order to provide a better insight into circular quaternion-valued random variables and to obtain tools for simulations, a procedure for generating circular quaternion-valued random variables is presented in this section. For the sake of simplicity the work is limited to the univariate case; however, the framework can easily be expanded for generating quaternion-valued random vectors.

Consider a quaternion-valued random variable of the form $Q = RU$ where R is a real-valued non-negative random variable and U is quaternion-valued random variable with uniform distribution on the surface of the unit hypersphere in $\mathbb{H}$ that is independent of R. The AQCF of Q can now be expressed as

$$\mathbf{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a) = E\left[e^{\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{q}^a}\right] = E\left[e^{\frac{\zeta}{4}r\mathbf{s}^{aH}\mathbf{u}^a}\right] = \int_0^\infty \oint_U e^{\frac{\zeta}{4}r\mathbf{s}^{aH}\mathbf{u}^a}\mathrm{d}F_{\boldsymbol{U}^a}(\mathbf{u}^a)\mathrm{d}F_R(r) \qquad (8.15)$$

where the symbol "$\oint_U$" denotes the integral on the unit hypersphere in $\mathbb{H}$, while $F_{\boldsymbol{U}^a}(\mathbf{u}^a)$ and $F_R(r)$ represent the cumulative distribution function (cdf) of U and R. Since R and U are independent, the expression in (8.15) yields

$$\mathbf{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a) = \int_0^\infty \mathbf{\Phi}_{\boldsymbol{U}^a}(r\mathbf{s}^a)\mathrm{d}F_R(r)$$

furthermore, U is a quaternion-valued circular random variable and therefore its AQCF takes the form $\mathbf{\Phi}_{\boldsymbol{U}^a}(r\mathbf{s}^a) = \phi(r|s|)$ where $\phi(\cdot) : \mathbb{R} \to \mathbb{R}$, resulting in

$$\mathbf{\Phi}_{\boldsymbol{Q}^a}(\mathbf{s}^a) = \int_0^\infty \phi(r|s|)\mathrm{d}F_R(r). \qquad (8.16)$$

Therefore, circular quaternion random variables with AQCF of the form in (8.16) are at hand. In addition, note that the AQCF of Q, as expressed in (8.16), is only a function of $|s|$ indicating that quaternion-valued circular random variables have circular symmetric AQCFs.

8.3.6 The $\mathbb{HC}$ interpretation and ζ-circularity

In some scenarios it becomes useful to consider a quaternion $q \in \mathbb{H}$ as the combination of two complex numbers whereby

$$q = \big(\underbrace{q_r + iq_i}_{z_1 \in \mathbb{C}}\big) + \big(\underbrace{q_j + iq_k}_{z_2 \in \mathbb{C}}\big)j$$

which is referred to as the Cayley-Dickson presentation [99, 102]. The joint augmented complex characteristic function of Z_1 and Z_2 is now given by

$$\boldsymbol{\Phi}_{\boldsymbol{Z}_1^a,\boldsymbol{Z}_2^a}(\mathbf{s}_1^a,\mathbf{s}_2^a) = E\left[e^{\left(\frac{i}{2}\left(\mathbf{s}_1^{aH}\mathbf{z}_1^a+\mathbf{s}_2^{aH}\mathbf{z}_2^a\right)\right)}\right]$$

where $\{\mathbf{z}_1^a,\mathbf{z}_1^a,\mathbf{s}_1^a,\mathbf{s}_1^a\} \in \mathbb{C}^2$. After some mathematical manipulations, the AQCF of $Q_{1\text{-}i} = Z_1 + Z_2 j$ can be formulated as

$$\boldsymbol{\Phi}_{\boldsymbol{Q}_{1\text{-}i}^a}(\mathbf{s}^a) = E\left[e^{\left(\frac{i}{4}\left(\mathbf{s}_{1\text{-}i}^{aH}\mathbf{q}_{1\text{-}i}^a\right)\right)}\right]$$

where $\mathbf{s}_{1\text{-}i}^a = \mathbf{B}[s_1, s_2, s_1^*, s_2^*]^T$, $\mathbf{q}_{1\text{-}i}^a = \mathbf{B}[z_1, z_2, z_1^*, z_2^*]^T$, and

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & 0 & j \\ 1 & 0 & 0 & -j \\ 0 & 1 & -j & 0 \\ 0 & 1 & j & 0 \end{bmatrix} \qquad \text{with} \qquad \frac{1}{2}\mathbf{B}^H = \mathbf{B}^{-1}.$$

Now consider a quaternion random variable such that $Q_{1\text{-}i} = Z_1 U_{1\text{-}j}$, where Z_1 is a random complex variable in the 1-i plane and $U_{1\text{-}j}$ is a uniform distribution on the unit circle in the 1-j plane. Note that since $U_{1\text{-}j}$ is uniformly distributed on the unit circle in the 1-j plane, the distribution of $Q_{1\text{-}i}$ will be invariant under rotations along the j-axis. After some mathematical manipulation the AQCF of $Q_{1\text{-}i}$ can be formulated as

$$\boldsymbol{\Phi}_{\boldsymbol{Q}_{1\text{-i}}^a}(\mathbf{s}^a) = E\left[e^{\left(\frac{i}{4}\left(\mathbf{s}_{1\text{-}i}^{aH}\mathbf{q}_{1\text{-}i}^a\right)\right)}\right] = E\left[e^{(i\xi)}\right] \tag{8.17}$$

where

$$\begin{aligned} \xi =& \Re\left(s_1^* q_1 + s_2^* q_2\right) \\ =& \Re\left(s_1^* z_1 u_{1\text{-}j_r} + s_2^* z_1 u_{1\text{-}j_j}\right) \\ =& \underbrace{\left(s_{1_r} z_{1_r} + s_{1_i} z_{1_i}\right)}_{\mathfrak{s}_r} u_{1\text{-}j_r} + \underbrace{\left(s_{2_r} z_{1_r} + s_{2_i} z_{1_i}\right)}_{\mathfrak{s}_j} u_{1\text{-}j_j} \end{aligned} \tag{8.18}$$

with $\mathfrak{s} = \mathfrak{s}_r + \mathfrak{s}_j j$. Furthermore, assuming Z_1 and $U_{1\text{-}j}$ are independent allows the AQCF of $Q_{1\text{-}i}$ as given in (8.17) to be rearranged into

$$\boldsymbol{\Phi}_{\boldsymbol{Q}_{1\text{-}i}^a}(\mathbf{s}^a) = E\left[e^{(i\Re(\mathfrak{s}^* u_{1\text{-}j}))}\right] = E\left[\boldsymbol{\Phi}_{\boldsymbol{U}_{1\text{-}j}^a}(\mathfrak{s}^a)\right] = E\left[\phi\left(|\mathfrak{s}|\right)\right]. \tag{8.19}$$

where $|\mathfrak{s}|$ can be expressed as

$$|\mathfrak{s}| = \sqrt{\Re^2\left(s_1 z_1^*\right) + \Re^2\left(s_2 z_1^*\right)}. \tag{8.20}$$

Therefore, if the quaternion random variable is j-circular its AQCF will take the form presented in (8.19). In addition, for such random variables, we have

$$\begin{aligned}\forall\theta \in [0, 2\pi) : \mathbf{\Phi}_{\boldsymbol{Q}^a_{1\text{-}i}}(\mathbf{s}^a) =& E\left[e^{\left(\zeta\Re(s^*q)\right)}\right] = E\left[e^{\left(\zeta\Re(s^*qe^{j\theta})\right)}\right] \\ =& E\left[e^{\left(\zeta\Re(sz_1u_{1\text{-}j}e^{j\theta})\right)}\right] = E\left[e^{\left(\zeta\Re(s^*e^{-j\theta}z_1q)\right)}\right] \\ =& E\left[e^{\left(\zeta\Re\left((se^{j\theta})^*\right)z_1q)\right)}\right] = \mathbf{\Phi}_{\boldsymbol{Q}^a_{1\text{-}i}}(\mathbf{s'}^a)\end{aligned}$$

with $s' = se^{j\theta}$. Therefore, $\mathbf{\Phi}_{\boldsymbol{Q}^a_{1\text{-}i}}(\mathbf{s}^a)$ is invariant under rotations along the j-axis[4]. Note that samples of an j-circular random variable can be generated by multiplying random samples of z_1 with $\cos(\theta) + j\sin(\theta)$ where $\theta \sim \mathcal{U}[0, 2\pi)$. Moreover, in this work, the quaternion domain is split into two complex pains along side the j-axis, similar results can be obtained by splitting the quaternion domain along side any of the imaginary axis or indeed along any line in the quaternion domain.

8.4 Quaternion-valued stable random variables

Considering the definition given in (8.1) and by extension, here we define a quaternion-valued random variable as α-stable if its real-valued components are jointly α-stable and therefore it is straightforward to show that such a quaternion-valued random variable also admits the stable property in (8.1). Furthermore, if a quaternion-valued random variable, $\boldsymbol{Q}$ is α-stable and has a symmetric elliptical distribution; then, from classical results in [138], the joint characteristic function of its real-valued components takes the form

$$\mathbf{\Phi}_{\boldsymbol{Q}}(\mathbf{s}_r, \mathbf{s}_i, \mathbf{s}_j, \mathbf{s}_k) = e^{-\left(\frac{1}{2}[\mathbf{s}_r,\mathbf{s}_i,\mathbf{s}_j,\mathbf{s}_k]\mathbf{\Sigma}[\mathbf{s}_r,\mathbf{s}_i,\mathbf{s}_j,\mathbf{s}_k]^T\right)^{\frac{\alpha}{2}}} \tag{8.21}$$

where $\mathbf{\Sigma}$ is a semi positive definite matrix dictating the elliptical shape of the distribution and hereafter is referred to as the covariation matrix. Note that for the case that $\alpha = 2$ the characteristic function in (8.21) is that of a multivariate Gaussian distribution with $\mathbf{\Sigma}$ denoting its covariance matrix. In addition, using the transformation in (2.8) it is trivial to show that the AQCF of quaternion-valued α-stable random variables with symmetric elliptical distributions takes the form

$$\mathbf{\Phi}_{\boldsymbol{Q}} = e^{-\left(\frac{1}{32}\mathbf{s}^{aH}\mathbf{\Sigma}_{\boldsymbol{Q}^a}\mathbf{s}^a\right)^{\frac{\alpha}{2}}} \tag{8.22}$$

where

$$\mathbf{\Sigma}_{\boldsymbol{Q}^a} = \frac{1}{16}\begin{bmatrix}\mathbf{I} & i\mathbf{I} & j\mathbf{I} & k\mathbf{I} \\ \mathbf{I} & i\mathbf{I} & -j\mathbf{I} & -k\mathbf{I} \\ \mathbf{I} & -i\mathbf{I} & j\mathbf{I} & -k\mathbf{I} \\ \mathbf{I} & -i\mathbf{I} & -j\mathbf{I} & k\mathbf{I}\end{bmatrix}^H \mathbf{\Sigma} \begin{bmatrix}\mathbf{I} & i\mathbf{I} & j\mathbf{I} & k\mathbf{I} \\ \mathbf{I} & i\mathbf{I} & -j\mathbf{I} & -k\mathbf{I} \\ \mathbf{I} & -i\mathbf{I} & j\mathbf{I} & -k\mathbf{I} \\ \mathbf{I} & -i\mathbf{I} & -j\mathbf{I} & k\mathbf{I}\end{bmatrix}.$$

[4]Note the same conclusion can be drawn from the expression in (8.20) and (8.19).

Once again, note that for the special case $\alpha = 2$ the AQCF given in (8.22) is that of a zero-mean quaternion-valued Gaussian distribution with the augmented covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{Q}^a}$.

Hereafter, in order to allow for mathematically tractable solutions, we focus on the case of quaternion-valued α-stable random variables with symmetric elliptical distributions, hereafter referred to as quaternion α-stable random variables for brevity. This class of stable random variables admit the following[5]:

- Let $\boldsymbol{G}$ be a zero-mean quaternion-valued Gaussian random vector with augmented covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{G}^a}$ and A a univariate real-valued positive random variable such that $\boldsymbol{\Phi}_A(s) = E[e^{sa}] = e^{-|s|^{\alpha/2}}$; then, the AQCF of the random variable $\boldsymbol{Q} = \sqrt{A}\boldsymbol{G}$ is given by

$$\begin{aligned} \Phi_{\boldsymbol{Q}^a}(\mathbf{s}^a) =& E\left[e^{\zeta \mathbf{s}^{aH}(\sqrt{a}\mathbf{q}^a)}|a\right] = E\left[E\left[e^{\eta \mathbf{s}^{aH}(\sqrt{a}\mathbf{q}^a)}|a\right]\right] \\ =& e^{-\left(\frac{1}{32}\boldsymbol{\Theta}_{\boldsymbol{Q}^a}\right)^{\alpha/2}} \text{ with } \boldsymbol{\Theta}_{\boldsymbol{Q}^a} = \mathbf{s}^{aH}\boldsymbol{\Sigma}_{\boldsymbol{G}^a}\mathbf{s}^a. \end{aligned} \tag{8.23}$$

  The expression in (8.23) implies that each quaternion α-stable random variable is composed of a zero-mean quaternion-valued Gaussian random variable that dictates its covariation matrix and a real-valued $\alpha/2$-stable random variable that dictates the α parameter. Furthermore, note that in the case where the underlying zero-man quaternion-valued Gaussian random variable is circular; then, $\boldsymbol{\Sigma}_{\mathbf{G}^a}$ is diagonal and the distribution of $\boldsymbol{Q}$ will be circular as well, while the real-valued components of $\boldsymbol{Q}$ are identically distributed, but not independent, which confirms the results in Section 8.3.4.

- Note that for a quaternion-valued α-stable random variable its AQCF in the formulation given in (8.22) is non-differentiable at $\mathbf{s}^a = 0$ for $\alpha \leq 1$ and only has finite first-order derivativeness at $\mathbf{s}^a = 0$ for $1 < \alpha \leq 2$; therefore, $\alpha = 2$ is the only case where such random variables have closed form second-order statistics that are obtainable through the framework set in Section 8.3.3. This comprises the performance of conventional adaptive filtering algorithms that are based on second-order statistics. This, issue is addressed in the next section.

8.5 Adaptive filtering

Consider the problem of finding the optimal widely-linear mapping that relates the quaternion-valued random variables $\mathbf{x}^a$ and $\mathbf{y}^a$, which can be formulated as

$$\mathbf{y}^a = \mathbf{W}^a\mathbf{x}^a. \tag{8.24}$$

[5]Most of the proofs have been omitted as they closely follow those of the real-valued vector case presented in [138].

In the case where $\{\mathbf{x}^a, \mathbf{y}^a\}$ are zero-mean and jointly Gaussian, multiplying both sides by $\mathbf{x}^{aH}$ yields

$$\mathbf{y}^a\mathbf{x}^{aH} = \mathbf{W}^a\mathbf{x}^a\mathbf{x}^{aH} \tag{8.25}$$

where the optimal solution in the MMSE sense referred to as the Wiener solution (see Chapter 2 and [71]) can now be obtained through taking the statistical expectation of the expression in (8.25) which gives

$$\mathbf{W}^a_{opt} = E\left[\mathbf{y}^a\mathbf{x}^{aH}\right]\left(E\left[\mathbf{x}^a\mathbf{x}^{aH}\right]\right)^{-1}. \tag{8.26}$$

However, the solution given in (8.26) is only optimal for the Gaussian case and cannot be generalized for the entire quaternion-valued α-stable random variables. In this section, the optimal filtering solution for the generality of quaternion-valued α-stable random variables is derived and then adaptive solutions are investigated.

8.5.1 The optimal filtering solution

Given the widely-linear mapping in (8.24) and considering $\mathbf{x}^a$ and $\mathbf{y}^a$ to be quaternion-valued α-stable random variables with symmetric elliptical distributions[6]; their joint AQCF can be expressed as

$$\mathbf{\Phi}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}(\mathbf{s}^a_y, \mathbf{s}^a_x) = E\left[e^{\frac{\varsigma}{4}\left(\mathbf{s}^{aH}_y\mathbf{y}^a + \mathbf{s}^{aH}_x\mathbf{x}^a\right)}\right] \tag{8.27}$$

where upon replacing $\mathbf{y}^a = \mathbf{W}^a\mathbf{x}^a$ in (8.27) we have

$$\mathbf{\Phi}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}(\mathbf{s}^a_y, \mathbf{s}^a_x) = E\left[e^{\frac{\varsigma}{4}\left(\mathbf{s}^{aH}_y\mathbf{W}^a\mathbf{x}^a + \mathbf{s}^{aH}_x\mathbf{x}^a\right)}\right] = E\left[e^{\frac{\varsigma}{4}\left(\left(\mathbf{s}^{aH}_y\mathbf{W}^a + \mathbf{s}^{aH}_x\right)\mathbf{x}^a\right)}\right]. \tag{8.28}$$

Now, considering the formulation in (8.22) for the AQCF of such random variables, the joint AQCF of $\mathbf{x}^a$ and $\mathbf{y}^a$ is given by

$$\mathbf{\Phi}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}(\mathbf{s}^a_y, \mathbf{s}^a_x) = \mathbf{\Phi}_X(\mathbf{s}^a_y\mathbf{W}^a + \mathbf{s}^a_x) = e^{-\left(\frac{1}{2}\mathbf{\Theta}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}\right)^{\frac{\alpha}{2}}} \tag{8.29}$$

where

$$\begin{aligned}
\mathbf{\Theta}_{\boldsymbol{Y}^a,\boldsymbol{X}^a} &= \left(\mathbf{s}^{aH}_y\mathbf{W}^a + \mathbf{s}^{aH}_x\right)\mathbf{\Sigma}_{\mathbf{X}^a}\left(\mathbf{W}^{aH}\mathbf{s}^a_y + \mathbf{s}^a_x\right) \\
&= \mathbf{s}^{aH}_y\underbrace{\mathbf{W}^a\mathbf{\Sigma}_{\mathbf{X}^a}\mathbf{W}^{aH}}_{\mathbf{\Sigma}_{\boldsymbol{Y}^a}}\mathbf{s}^a_y + \mathbf{s}^{aH}_y\underbrace{\mathbf{W}^a\mathbf{\Sigma}_{\boldsymbol{X}^a}}_{\mathbf{\Sigma}_{\boldsymbol{Y}^a\boldsymbol{X}^a}}\mathbf{s}^a_x + \mathbf{s}^{aH}_x\underbrace{\mathbf{\Sigma}_{\mathbf{X}^a}\mathbf{W}^{aH}}_{\mathbf{\Sigma}_{\boldsymbol{X}^a\boldsymbol{Y}^a}}\mathbf{s}^a_y + \mathbf{s}^{aH}_x\mathbf{\Sigma}_{\boldsymbol{X}^a}\mathbf{s}^{aH}_x. \\
&= \begin{bmatrix}\mathbf{s}^{aH}_y & \mathbf{s}^{aH}_x\end{bmatrix}\underbrace{\begin{bmatrix}\mathbf{\Sigma}_{\boldsymbol{Y}^a} & \mathbf{\Sigma}_{\boldsymbol{Y}^a\boldsymbol{X}^a} \\ \mathbf{\Sigma}_{\boldsymbol{X}^a\boldsymbol{Y}^a} & \mathbf{\Sigma}_{\boldsymbol{X}^a}\end{bmatrix}}_{\mathbf{\Sigma}_{\boldsymbol{Y}^a,\mathbf{X}^a}}\begin{bmatrix}\mathbf{s}^a_y \\ \mathbf{s}^a_x\end{bmatrix}.
\end{aligned} \tag{8.30}$$

[6]In incidences where the filtering of α-stable random variables is considered it is implicitly implied that $\alpha > 1$, so that conditional exceptions $E[y|x]$ exist and are finite.

From the expression in (8.30), note that $\mathbf{\Sigma}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}$ is in fact the Hessian matrix of $\mathbf{\Theta}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}$; therefore, the optimal filtering solution can be expressed as

$$\mathbf{W}^a_{opt} = \mathbf{\Sigma}_{\boldsymbol{X}^a\boldsymbol{Y}^a}\left(\mathbf{\Sigma}_{\boldsymbol{X}^a}\right)^{-1}$$

where $\mathbf{\Sigma}_{\boldsymbol{X}^a\boldsymbol{Y}^a}$ and $\mathbf{\Sigma}_{\boldsymbol{X}^a}$ are elements of the Hessian matrix of

$$\mathbf{\Theta}_{\boldsymbol{Y}^a,\boldsymbol{X}^a} = \Bigg(-32\ln\left(\mathbf{\Phi}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}(\mathbf{s}^a_y,\mathbf{s}^a_x)\right)\Bigg)^{\frac{2}{\alpha}}.$$

In most cases, however, applying block based estimators for finding the optimal filtering solution is computationally complex and rather inadequate when dealing with non-stationary signals. Thus, adaptive approaches are required.

8.5.2 Steepest-descent algorithms

Although $\mathbf{\Theta}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}$ can be evaluated from empirical estimation of $\mathbf{\Phi}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}$ using samples of the signal; however, obtaining the Hessian matrix of $\mathbf{\Theta}_{\boldsymbol{Y}^a,\boldsymbol{X}^a}$ through numerical methods is computationally inefficient. Moreover, applying block based estimators is rather inadequate specially when dealing with non-stationary signals. Therefore, following the approaches in [142–144], an adaptive approach is next investigated. To this end, consider the widely-linear mapping in (8.24) in its adaptive formulation, given by

$$\hat{y}_n = \mathbf{h}^T_n\mathbf{x}_n + \mathbf{g}^T_n\mathbf{x}^i_n + \mathbf{u}^T_n\mathbf{x}^j_n + \mathbf{v}^T_n\mathbf{x}^k_n = \mathbf{w}^{aT}_n\mathbf{x}^a_n$$

the coefficients of which are updated at each time instant in a steepest-descent fashion minimizing the cost function $J = E\left[|y-\hat{y}|^p\right] = E\left[|\varepsilon|^p\right]$ in its instantaneous form given by

$$J_n = |\varepsilon_n|^P = (\varepsilon_n\varepsilon^*_n)^{\frac{p}{2}}$$

where $\varepsilon_n = y_n - \hat{y}_n$ and $1 < p < \alpha$. The gradients of J_n can be calculated through the $\mathbb{HR}$-calculus and is given by

$$\nabla_{\mathbf{w}^{a*}_n} J_n = \nabla_{\mathbf{w}^{a*}_n}\left(\epsilon_n\epsilon^*_n\right)^{\frac{p}{2}} = -\frac{3}{4}\frac{p}{2}|\epsilon_n|^{p-2}\epsilon_n\mathbf{x}^{a*}_n.$$

Therefore the update of the weight vector can be expressed as

$$\mathbf{w}^a_{n+1} = \mathbf{w}^a_n + \mu|\varepsilon_n|^{p-2}\varepsilon_n\mathbf{x}^{a*}_n \tag{8.31}$$

where $\mu \in \mathbb{R}^+$ represents an adaptation gain.

Considering the wight error given by $\boldsymbol{\epsilon}^a_n = \mathbf{w}^a_{opt} - \mathbf{w}^a_n$ and the update equation in (8.31), the evolution of the weight vector error can be expressed as

$$\boldsymbol{\epsilon}^a_{n+1} = \boldsymbol{\epsilon}^a_n - \mu|\varepsilon_n|^{p-2}\varepsilon_n\mathbf{x}^{a*}_n$$

where replacing $\varepsilon_n = \hat{y}_n - y_n = \boldsymbol{\epsilon}_n^{aT}\mathbf{x}_n^a$ and after some mathematical manipulations we have

$$\boldsymbol{\epsilon}_{n+1}^a = \boldsymbol{\epsilon}_n^a\left(\mathbf{I} - \mu|\varepsilon_n|^{p-2}\mathbf{x}_n^a\mathbf{x}_n^{aH}\right). \tag{8.32}$$

The expression in (8.32) gives the mapping of the weight vector error from one time instant to the next; therefore, for the weight vector to converge the transform has to be a contracting one, which gives a bound for the adaptation gain as

$$\mu \leq \frac{1}{|\varepsilon_n|^{p-2}\mathrm{Tr}(\mathbf{x}_n^a\mathbf{x}_n^{aH})}. \tag{8.33}$$

Note that for the special case where $\alpha = 2$ and $p = 2$, the update equation in (8.31) simplifies to that of the WL-QLMS; in addition, selecting $\mu = \mathrm{Tr}(\mathbf{x}_n^a\mathbf{x}_n^{aH})$ results in the normalized WL-QLMS algorithm. However, regarding the general case where $1 < \alpha \leq 2$ the following statements can be made:

- The condition $p > 1$ ensures the cost function $J_n = |\varepsilon_n|^p$ is convex and differentiable for all values of ε_n.
- If the input and output of the filter are jointly α-stable; then, ε is a quaternion-valued α-stable variable with α-stable real-valued components. Furthermore,

  $$J = E\left[|\varepsilon|^p\right] = E\left[\left(\varepsilon_r^2 + \varepsilon_i^2 + \varepsilon_j^2 + \varepsilon_k^2\right)^{\frac{p}{2}}\right] \leq E\left[|\varepsilon_r|^p\right] + E\left[|\varepsilon_i|^p\right] + E\left[|\varepsilon_j|^p\right] + E\left[|\varepsilon_k|^p\right].$$

  Therefore, the condition $p \leq \alpha$ will ensure that $E\left[|\varepsilon_r|^p\right]$, $E\left[|\varepsilon_i|^p\right]$, $E\left[|\varepsilon_j|^p\right]$, and $E\left[|\varepsilon_k|^p\right]$ exist and hence the cost function is bounded.
- Since $\varepsilon_n = \boldsymbol{\epsilon}_n^{aT}\mathbf{x}_n^a$, for a given weight vector error at time instant n, sharp increases in $\|\mathbf{x}_n^a\|_2$ will result in a rise in $|\varepsilon_n|$ and a fall in $|\varepsilon_n|^{p-2}$; thus, the presence of the term $|\varepsilon_n|^{p-2}$ in the weight vector update equation in (8.31) helps to regulate the adaptation step-size when dealing with heavy tails. This is shown in Figure 8.4, where the WL-QLMS, normalized WL-QLMS, and a steepest descent algorithm with $p = 1.1$ are implemented to estimate the weight vector of an autoregressive process of length 4 that is driven by a proper quaternion-valued 1.9-stable noise.

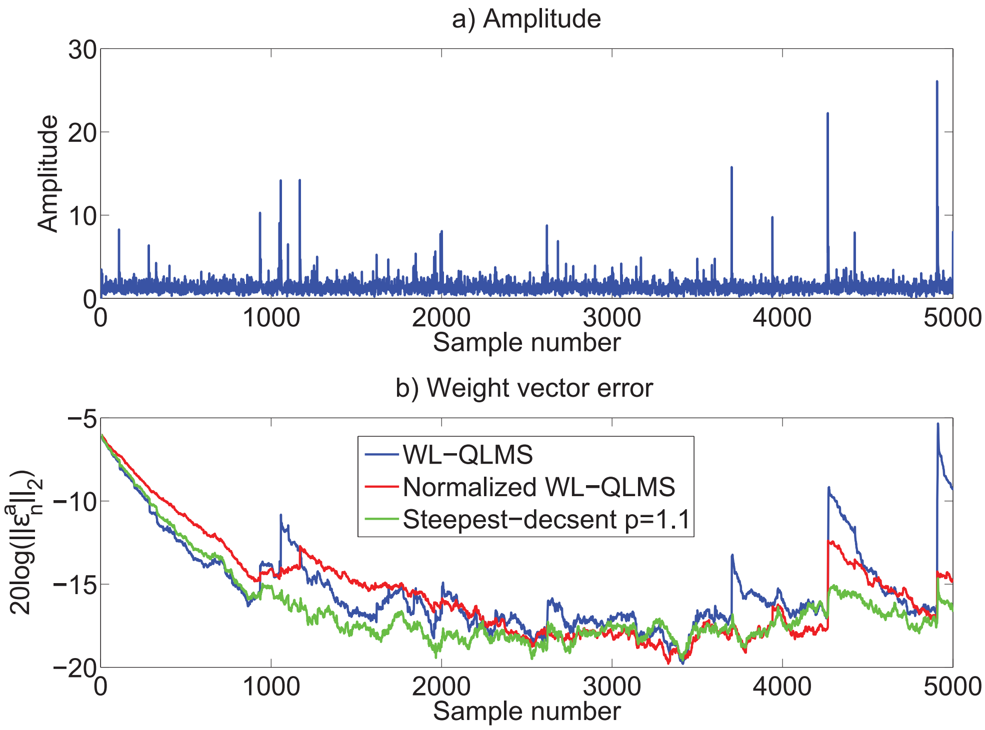

Figure 8.4: Estimation of the weight vector of an autoregressive process of length 4 that is driven by a proper quaternion-valued 1.9-stable noise: a) amplitude of the autoregressive process, b) weight vector estimation error.

8.5.3 The augmented quaternion particle filter

Although the steepest-decent approach is a computationally effective method for processing α-stable random signals; however, lack of convergence bounds, limits their use. In order to present an inclusive framework for processing α-stable random signals the augmented quaternion particle filter is next derived. To this end, consider the general widely-linear state space model given by

$$\begin{aligned}\mathbf{x}_n^a =&\mathbf{A}_n^a\mathbf{x}_{n-1}^a + \boldsymbol{\nu}_n^a\\ \mathbf{y}_n^a =&\mathbf{H}_n^a\mathbf{x}_n^a + \boldsymbol{\omega}_n^a\end{aligned}$$

where $\mathbf{x}_n^a$ and $\mathbf{y}_n^a$ are the augmented state and observations at time instant n, while $\boldsymbol{\nu}_n^a$ and $\boldsymbol{\omega}_n^a$ represent the state evolution and observation noise.

Taking the conventional particle filtering approach [147, 148], the AQCF of the state vector sequence $\mathbf{x}_{0:n}^a = \{\mathbf{x}_0^a, \dots, \mathbf{x}_n^a\}$ conditional on the observation sequence $\mathbf{y}_{1:n}^a =$

$\{\mathbf{y}_1^a, \ldots, \mathbf{y}_n^a\}$ can be expressed as

$$\begin{aligned}\mathbf{\Phi}_{\boldsymbol{X}_{0:n}^a|\boldsymbol{Y}_{1:n}^a}(\mathbf{s}^a) =& \int_{\mathcal{D}_{\boldsymbol{X}_{0:n}^a}} e^{\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{x}_{0:n}^a} P_{\boldsymbol{X}_{0:n}^a}(\mathbf{x}_{0:n}^a|\mathbf{y}_{1:n}^a)\mathrm{d}\mathbf{x}_{0:n}^a \\ \approx& \frac{1}{\sum_{m=1}^M w_n^{\{m\}}} \sum_{m=1}^M w_n^{\{m\}} e^{\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{x}_{0:n}^{a\{m\}}}\end{aligned}$$

where $P_{\boldsymbol{X}_{0:n}^a}(\mathbf{x}_{0:n}^a|\mathbf{y}_{1:n}^a)$ denotes the probability of the augmented state vector sequence $\mathbf{x}_{0:n}^a$, conditional to the augmented observation sequence $\mathbf{y}_{1:n}^a$, whereas $\mathcal{D}_{\boldsymbol{X}_{0:n}^a}$ denotes the domain of $\boldsymbol{X}_{0:n}^a$, while $\mathbf{x}_{0:n}^{a\{m\}}$ and $w_n^{\{m\}}$ are independent particles drawn from the distribution of $\boldsymbol{X}_{0:n}^a$, or its importance function $\mathcal{P}(\mathbf{x}_{0:n}^a|\mathbf{y}_{1:n}^a)$, and their associated weights given by

$$w_n^{\{m\}} \propto \frac{P_{\boldsymbol{X}_{0:n}^a}(\mathbf{x}_{0:n}^a) P_{\boldsymbol{Y}_{1:n}^a|\boldsymbol{X}_{0:n}^a}(\mathbf{y}_{1:n}^a|\mathbf{x}_{0:n}^a)}{\mathcal{P}(\mathbf{x}_{0:n}^a|\mathbf{y}_{1:n}^a)}. \tag{8.34}$$

Assuming that the current state is independent from future observations and that the importance function is selected to be factorisable so that

$$\mathcal{P}(\mathbf{x}_{0:n+1}^a|\mathbf{y}_{1:n+1}^a) = \mathcal{P}(\mathbf{x}_{0:n}^a|\mathbf{y}_{1:n}^a)\mathcal{P}(\mathbf{x}_{n+1}^a|\mathbf{x}_{0:n}^a, \mathbf{y}_{1:n+1}^a)$$

allows the weights to be updated sequentially as

$$w_{n+1}^{\{m\}} \propto \frac{P_{\boldsymbol{Y}_{col,n+1}^a|\boldsymbol{X}_{n+1}^a}(\mathbf{y}_{col,n+1}^a|\mathbf{x}_{n+1}^a) P_{\boldsymbol{X}_{n+1}^a|\boldsymbol{X}_n^a}(\mathbf{x}_{n+1}^a|\mathbf{x}_n^a)}{\mathcal{P}(\mathbf{x}_{n+1}^a|\mathbf{x}_{0:n}^a, \mathbf{y}_{1:n+1}^a)} w_n^{\{m\}} \tag{8.35}$$

where distribution $P_{\boldsymbol{X}_{n+1}^a|\boldsymbol{X}_n^a}(\mathbf{x}_{n+1}^a|\mathbf{x}_n^a)$ is determined by the state evolution function.

Furthermore, if the distribution of $\boldsymbol{X}_{0:n}^a$ is approximated to be elliptically contoured, it can be fully described through the mean estimate given by

$$E[\mathbf{x}_{0:n}^a] \approx \frac{1}{\sum_{m=1}^M w_n^{\{m\}}} \sum_{m=1}^M w_n^{\{m\}} \mathbf{x}_{0:n}^{a\{m\}} \tag{8.36}$$

and covariation matrix estimate that is calculable from the widely-linear regression

$$\mathbf{s}^{aH}\mathbf{\Sigma}_{\boldsymbol{X}_{0:n}^a}\mathbf{s}^a = \left(-32\ln\left(\hat{\mathbf{\Phi}}_{\boldsymbol{X}_{0:n}^a}(\mathbf{s}^a)\right)\right)^{\frac{2}{\alpha}} \tag{8.37}$$

where

$$\hat{\mathbf{\Phi}}_{\boldsymbol{X}_{0:n}^a}(\mathbf{s}^a) = \frac{1}{\sum_{m=1}^M w_n^{\{m\}}} \sum_{m=1}^M w_n^{\{m\}} e^{\frac{\zeta}{4}\left(\mathbf{s}^{aH}\left(\mathbf{x}_{0:n}^{a\{m\}} - E[\mathbf{x}_{0:n}^a]\right)\right)}.$$

In addition, it can be shown that the AQCF of the state vector at two consecutive time instances are related according to

$$\begin{aligned}\mathbf{\Phi}_{\boldsymbol{X}_{n+1}^a}(\mathbf{s}^a) =& E\left[e^{\frac{\zeta}{4}\left(\mathbf{s}^{aH}\mathbf{A}_n^a\mathbf{x}_n^a + \mathbf{s}^{aH}\boldsymbol{\nu}_n^a\right)}\right] = \mathbf{\Phi}_{\boldsymbol{X}_n^a}(\mathbf{A}_n^{aH}\mathbf{s}^a)\mathbf{\Phi}_{\boldsymbol{\nu}_n^a} \\ =& e^{\left(\frac{\zeta}{4}\mathbf{s}^{aH}\mathbf{A}_n^a E[\mathbf{x}_n^a]\right)} e^{-\left(\frac{1}{32}\mathbf{s}^{aH}\mathbf{A}_n^a\mathbf{\Sigma}_{\boldsymbol{X}_n^a}\mathbf{A}_n^{aH}\mathbf{s}^a\right)} e^{-\left(\frac{1}{32}\mathbf{s}^{aH}\mathbf{\Sigma}_{\boldsymbol{\nu}_n^a}\mathbf{s}^a\right)}.\end{aligned}$$

Therefore, the covariation matrix of $\boldsymbol{X}_n^a$ can be found through the widely-linear regression

$$\mathbf{s}^{aH}\boldsymbol{\Sigma}_{\boldsymbol{X}_{n+1}^a}\mathbf{s}^a = \left(\left(\mathbf{s}^{aH}\mathbf{A}_n^a\boldsymbol{\Sigma}_{\boldsymbol{X}_n^a}\mathbf{A}_n^{aH}\mathbf{s}^a\right)^{\alpha/2} + \left(\mathbf{s}^{aH}\boldsymbol{\Sigma}_{\boldsymbol{\nu}_n^a}\mathbf{s}^a\right)^{\alpha/2}\right)^{2/\alpha}$$

while $E[\mathbf{x}_{n+1}^a] = \mathbf{A}_n^a E[\mathbf{x}_n^a]$. This allows to propagate the state vector statistics without the need to manipulate large number of particles. The operations of such an particle filtering algorithm are summarized in Algorithm 13.

Algorithm 13. Augmented Quaternion Particle Filter (AQPF)

Initialize:

Draw samples $\mathbf{x}_0^{a\{m\}}$ and assign weights $w_0^{\{m\}}$ using $\mathcal{P}(\mathbf{x}_0^a)$.

At each time instant:

1. Sample from the importance density $\mathcal{P}(\mathbf{x}_{0:n-1}^a|\mathbf{y}_{1:n-1}^a)$ and assign weights through (8.34).
2. Track samples through the state evaluation function.
3. Reassign weights through (8.35).
4. Approximate the distribution of $\boldsymbol{X}_{0:n}^a$ with that of an elliptically contoured α-stable distribution with mean and covariation matrix given in (8.36) and (8.37).
5. Draw particles from a quaternion-valued elliptically contoured α-stable distribution with the mean and covariation matrix calculated in the previous step to be propagated to the next stage.

As an example, consider the quaternion-valued dynamic system characterized in discrete time through the state space equations

$$\begin{aligned}\mathbf{x}_n &= \begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}\mathbf{x}_{n-1} + \begin{bmatrix}0.5\\ 1\end{bmatrix}\nu_n\\ y_n &= \begin{bmatrix}1 & 0\end{bmatrix}\mathbf{x}_n + \omega_n\end{aligned}$$

where ν_n is a zero-mean unit variance second-order proper Gaussian noise, while ω_n is a zero-mean elliptically distributed 1.5-stable noise with the covariation matrix

$$\boldsymbol{\Sigma}_{\boldsymbol{\omega}^a} = \begin{bmatrix}13 & 3 & 3 & 3\\ 3 & 13 & 3 & 3\\ 3 & 3 & 13 & 3\\ 3 & 3 & 3 & 13\end{bmatrix}.$$

The AQPF was implemented in order to track the state vector. The estimates of the state vector element that is observed through the observation function is shown in Figure 8.5, while the estimates of the state vector element that is not observable are shown in Figure 8.6. Note that despite the presence of a heavy-tailed 1.5-stable observation noise, the AQPF was able to track the state vector.

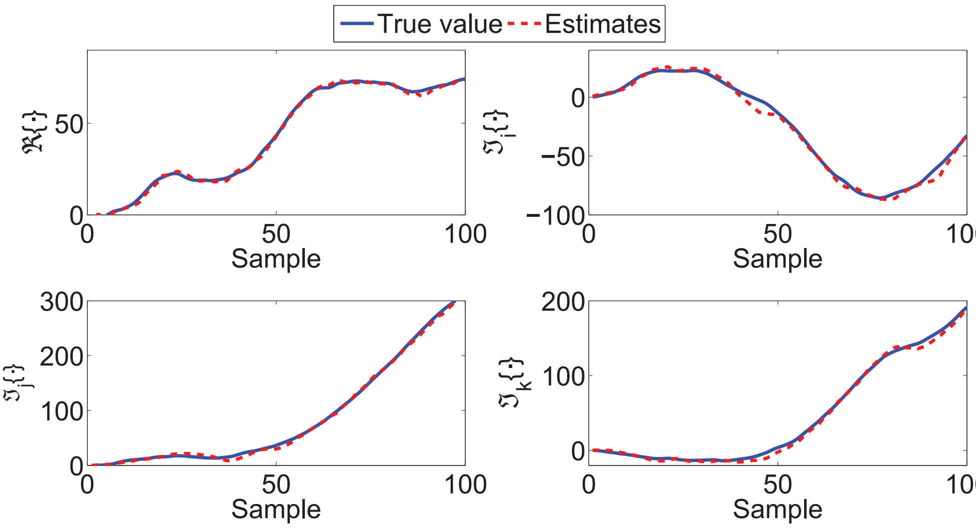

Figure 8.5: State vector estimation in presence of 1.5-stable observation noise. The estimates of the state vector element that is observed through the observation function are shown.

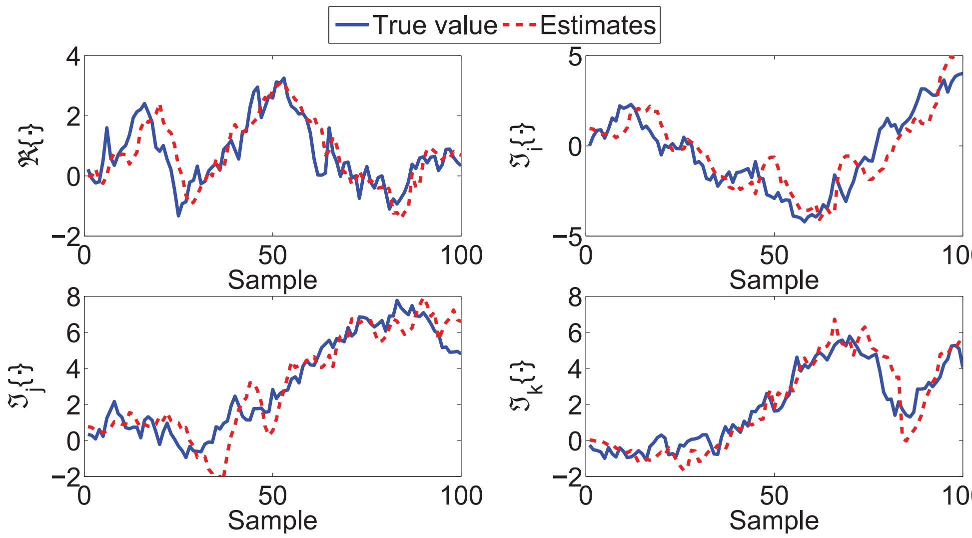

Figure 8.6: State vector estimation in presence of 1.5-stable observation noise. The estimates of the state vector element that is not observable through the observation function are shown.

8.6 Conclusion

The augmented quaternion characteristic function was revised and its properties from a geometric point of view and its link to statistical moments were investigated. In addition, adaptive filtering of quaternion-valued α-stable random variables with symmetric elliptical distributions was considered. The contributions of this chapter can be summarized as follows:

- The link between the AQCF and different types of circularity was investigated establishing a structure for the AQCF of circular and ζ-circular quaternion-valued random variables and a framework for sample generation. furthermore, it was established that for the case of quaternion-valued random variables with i.i.d components, circularity happens only in the Gaussian case.
- The statistical manipulation of quaternion-valued α-stable random variables was considered, where an optimal filtering solution was derived through the use of the AQCF. In addition, adaptive filtering of quaternion-valued α-stable random signals was investigated, where the augmented quaternion particle filter was introduced for sequential state estimation of quaternion-valued α-stable signals.

Chapter 9

Conclusion

9.1 Conclusion and contributions

Motivated by the recent developments in quaternion-valued signal processing and their natural ability to model three-dimensional rotations; in this thesis, quaternion-valued signal processing algorithms dealing with the notion of phase and frequency in the quaternion domain with practical applications in smart grids, target tracking, and tracking the rotations of objects in the three-dimensional space were developed. The use of quaternions in these applications where the underlying signal is three-dimensional by nature, has allowed for the derivation of signal processing algorithms that in addition to having a rigorous physical interpretation, can account for all the information in the signal and outperform their complex or real-valued counterparts. The contributions of this thesis are summarized in the following:

1. Although widely-linear adaptive filters are optimal for the generality of quaternion-valued signals, they can be simplified into strictly linear adaptive filters in cases where the signal is proper, significantly reducing computational complexity. To this end, a novel real-time tracker of quaternion impropriety was developed which allows to identify the degree of impropriety of a signal in real-time, so that the instances when a non-stationary signal changes its statistic can be identified and an estimator that best suits the signal can be selected. The work includes comprehensive theoretical analysis of the mean and mean-square behavior of the impropriety tracker, which are verified through simulations.
2. Adaptive phase-only estimation of quaternion-valued signals was considered and a class of quaternion phase estimators was developed. This was achieved through updating the weights of the adaptive filter at each time instant according to a cost function of the phase error in a steepest-descent fashion. The performance of the algorithms was analyzed, and a geometrical interpretation of the operations of

the proposed algorithms was presented. The developed phase-only estimator was used in a number of practical applications for estimating rotations in the three-dimensional space, where it showed superior performance as compared to that of real-valued quadrivariate algorithms based on the application of rotation matrices.

3. A widely-linear complex-valued frequency estimator for three-phase power systems that can outperform its complex-valued counterparts in addition to having consistent performance under both nominal and fault conditions was developed. For rigor, the performance of the developed frequency estimator was quantified and compared to that of its counterparts. Finally, the analysis was confirmed through simulations on both synthetic data and real-world data recordings, where the developed frequency estimator showed outstanding performance.

4. Quaternions were used for frequency estimation in three-phase power systems, where the multidimensional nature of quaternions allowed a state space model for the three-phase power signal to be developed that incorporates voltage measurements from all phases of the power system without the need for using the Clarke transform. In addition, the state space model was designed such that its elements could be applied for the estimation of the system voltage phasors. The performance of the developed frequency estimator was quantified and compared to that of traditional frequency estimators established on complex-valued algorithms through simulations, where it was shown to achieve a lower steady-state mean square error.

5. A fully distributed sequential state estimator for quaternion-valued signals was developed. This was achieved through decomposing the operations of the centralized quaternion Kalman filter in such a fashion that they can be performed in a distributed manner. The performance of the developed algorithm was analyzed establishing a recursive expression for the estimation error. The developed algorithm was implemented for estimating the fundamental frequency of three-phase power distribution networks and for collaborative target tracking.

6. Having established that quaternion-valued random variables with i.i.d components, circularity happens only in the Gaussian case, a structure for the augmented quaternion characteristic function of circular and ζ-circular quaternion-valued random variables and a framework for generating samples from such distributions was established. In addition, statistics of quaternion-valued α-stable random variable was considered, leading to an expression for the optimal filtering solution when dealing with zero-mean elliptically contoured α-stable random signals with $\alpha > 1$. Furthermore, adaptive filtering of quaternion-valued α-stable random variables was investigated and the augmented quaternion particle filter was introduced for adaptive filtering of non-Gaussian quaternion-valued signals.

9.2 Future work

The research into quaternion-valued signal processing has brought to light a number of new directions that are worth pursuing. These directions are as follows:

1. *Fly-by-Wire Systems*: The introduction of automated flight control systems has made possible aircraft designs that sacrifice aerodynamic stability in order to incorporate stealth technology into their shape and/or operate more efficiently. One of the most important tasks of such systems is to track and control the pitch, roll, and yaw angles of the aircraft in real-time. Given the natural ability of quaternions to model rotations, the opportunity arises to develop a rigorous integrated quaternion-value flight control system.

2. *Power System Analysis*: For more than 50 years three-phase power systems have been analyzed through complex-valued mathematical techniques; however, this is a compromised approach as complex numbers lack the dimensionality necessary to represent three-phase power signals. Therefore, the opportunity arises to develop a quaternion-valued framework for modeling the power grid and analyzing its performance.

3. *Distributed Filtering of α-Stable Random Signals*: The particle filter is the only adaptive filtering algorithm applicable to the generality of α-stable random signals; however, its computational burden remains as its major drawback when it comes to real-time implementation. Given the popularity that sensor networks and distributed signal processing has gained, distributing the computational load of the particle filter among agents of a network can be seen as a practical technique for addressing the computational complexity of particle filters.

Bibliography

[1] S. Haykin, *Adaptive filter theory.* Prentice Hall, 1996.

[2] P. S. R. Diniz, *Adaptive filtering: Algorithms and practical implementation.* Springer, 2008.

[3] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and adaptive signal processing: Spectral estimation, signal modeling, adaptive filtering, and array processing.* Artech House, 2005.

[4] R. L. Plackett, "Some theorems in least squares," *Biometrika*, vol. 37, no. 1/2, pp. 149–157, Jun 1950.

[5] L. Ljung, M. Morf, and D. Falconer, "Fast calculation of gain matrices for recursive estimation schemes," *International Journal of Control*, vol. 27, no. 1, pp. 1–19, 1978.

[6] M. Morf, T. Kailath, and L. Ljung, "Fast algorithms for recursive identification," *In Proceedings of IEEE Conference on Decision and Control including the 15th Symposium on Adaptive Processes*, pp. 916–921, December 1976.

[7] S. Qiao, "Fast adaptive RLS algorithms: A generalized inverse approach and analysis," *IEEE Transactions on Signal Processing*, vol. 39, no. 6, pp. 1455–1459, June 1991.

[8] D. T. M. Slock, L. Chisci, H. Lev-Ari, and T. Kailath,, "Modular and numerically stable fast transversal filters for multichannel and multiexperiment RLS," *IEEE transacions on Signal Processing*, vol. 40, no. 4, pp. 784–802, April 1992.

[9] B. Widrow and M. E. Hoff Jr, "Adaptive switching circuits," *IRE WESCON Convention Record, Part 4*, pp. 96–104, August 1960.

[10] B. Widrow and S. D. Stearns, *Adaptive signal processing.* Prentice Hall, 1985.

[11] W. P. Ang and B. Farhang-Boroujeny, "A new class of gradient adaptive step-size LMS algorithms," *IEEE Transactions on Signal Processing*, vol. 49, no. 4, April 2001.

[12] J. Nagumo and A. Noda, "A learning method for system identification," *IEEE Transactions on Automatic Control*, vol. 12, no. 3, pp. 282–287, June 1967.

[13] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 115–118, February 2004.

[14] S. C. Douglas, "Generalized gradient adaptive step sizes for stochastic gradient adaptive filters," *In Proceesings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 1396–1399, 1995.

[15] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME Journal of Basic Engineering*, vol. 82, Series D, pp. 35–45, 1960.

[16] M. S. Grewal and A. P. Andrews, "Applications of Kalman filtering in aerospace 1960 to the present [historical perspectives]," *IEEE Control systems*, vol. 30, no. 3, pp. 69–78, June 2010.

[17] S. Schmidt, "Application of state-space methods to navigation problems," *Advanced Control Systems*, vol. 3, pp. 293–340, 1966.

[18] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[19] M. S. Grewal and A. P. Andrews, *Kalman filtering: Theory and practice using MATLAB*. Wiley, 2015.

[20] Y. M. Tang and J. A. Borrie, "Missile guidance based on Kalman filter estimation of target maneuver," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-20, no. 6, pp. 736–741, November 1984.

[21] N. F. Palumbo, G. A. Harrison, R. A. Blauwkamp, and J. K. Marquart, "Guidance filter fundamentals," *Johns Hopkins APL Technical Digest*, vol. 29, no. 1, 2010.

[22] D. Lautier, A. Javaheri, and A. Galli, "Filtering in finance," *Wilmott Magazine*, no. 5, pp. 67–83, 2003.

[23] A. Sinha, T. Kirubarajan, and Y. Bar-Shalom, "Application of the Kalman-Levy filter for tracking maneuvering targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 1099–1107, July 2007.

[24] S. Y. Chen, "Kalman filter for robot vision: A survey," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 11, pp. 4409–4420, November 2012.

[25] T. Y. Al-Naffouri, "An EM-based forward-backward Kalman filter for the estimation of time-variant channels in OFDM," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3924–3930, July 2007.

[26] S. Liu and Z. Tian, "A Kalman-PDA approach to soft-decision equalization for frequency-selective MIMO channels," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3819–3830, October 2005.

[27] S. Haykin, Ed., *Kalman filtering and neural networks.* John Wiley & Sons, 2001.

[28] H. Hefies, "The effect of erroneous models on the Kalm filter response," *IEEE Transactions on Automated Control*, vol. AC-11, no. 3, pp. 541–543, July 1966.

[29] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering," *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 175–184, April 1970.

[30] L. W. Nelson and E. Stear, "The simultaneous on-line estimation of parameters and states in linear systems," *IEEE Transactions on Automatic Control*, vol. 21, no. 1, pp. 94–98, February 1976.

[31] E. A. Wan and A. T. Nelson, "Dual Kalman filtering methods for nonlinear prediction, smoothing and estimation," in *Advances in Neural Information Processing Systems 9*, M. Mozer, M. Jordan, and T. Petsche, Eds. MIT Press, 1997, pp. 793–799.

[32] A. T. Nelson, "Nonlinear estimation and modeling of noisy time series by dual *Kalman* filtering methods," Ph.D. dissertation, Oregon Graduate Institute of Science and Technology, 2000.

[33] D. P. Mandic and V. S. L. Goh, *Complex valued nonlinear adaptive filters: Noncircularity, widely linear and neural models.* Wielly, 2009.

[34] R. P. Feynmn, R. B. Leighton, and M. Sands, *The Feynman lectures on physics.* Basic Books, 2013, vol. 1 and 2, originally published in 1964.

[35] A. C. Phillips, *Introduction to quantum mechanics.* Wielly, 2003.

[36] R. P. Feynmn, R. B. Leighton, and M. Sands, *The Feynman lectures on physics.* Addison-Wesley, 2006, vol. 3, originally published in 1965.

[37] A. B. Carlson, P. B. Crilly, and J. C. Rutledge, *Communication systems: An Introduction to signals and noise in electrical communications*, 4th ed. McGraw-Hill, 2002.

[38] C. A. Desoer and E. H. Kuh, *Basic circuit theory.* McGraw-Hill, 2009, vol. 2.

[39] B. Widrow, J. McCool, and M. Ball, "The complex LMS algorithm," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 719–720, April 1975.

[40] B. Farhang-Boroujeny, *Adaptive filters: Theory and applications.* Wiley, 1998.

[41] Y. Bar-Ness and F. Haber, "Interference canceler array with reference generating loop," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-21, no. 5, pp. 654–662, September 1985.

[42] S. S. Narayan, "Frequency domain least-mean-square algorithm," *Proceedings of the IEEE*, vol. 69, no. 1, pp. 124–126, January 1981.

[43] P. J. Schreier and L. L. Scharf, "Second-order analysis of improper complex random vectors and processes," *IEEE Transactions on Signal Processing*, vol. 51, no. 3, pp. 714–725, March 2003.

[44] T. Adali, P. J. Schreier, and L. L. Scharf, "Complex-valued signal processing: The proper way to deal with impropriety," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5101–5125, November 2011.

[45] S. Javidi, M. Pedzisz, S. L. Goh, and D. P. Mandic, "The augmented complex least mean square algorithm," *In Proceedings of The 1st IARP Workshop on Cognitive Information Processing*, pp. 54–57, 2008.

[46] D. P. Mandic, S. Javidi, V. S. L. Goh, A. Kuh, and K. Aihara, "Complex-valued prediction of wind profile using augmented complex statistics," *Renewable Energy*, vol. 43, no. 1, pp. 196–201, 2007.

[47] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *IEE Proceedings: Communications, Radar, and Signal Processing*, vol. 130, no. 1, pp. 11–16, February 1983.

[48] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Mathematische Annalen*, vol. 97, pp. 357–375, December 1927, published in German.

[49] F. D. Neeser and J. L. Massey, "Proper complex random processes with applications to information theory," *IEEE Transactions on Information Theory*, vol. 39, no. 4, pp. 1293–1302, July 1993.

[50] V. S. L. Goh and D. P. Mandic, "An augmented extended Kalman filter algorithm for complex-valued recurrent neural networks," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, pp. V561–V564, May 2006.

[51] D. H. Dini and D. P. Mandic, "Class of widely linear complex Kalman filters," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 5, pp. 775–786, May 2012.

[52] D. H. Dini, D. P. Mandic, and S. J. Julier, "A widely linear complex unscented Kalman filter," *IEEE Signal Processing Letters*, vol. 18, no. 11, pp. 623–626, November 2011.

[53] E. Clarke, *Circuit analysis of AC power systems.* John Wiley & Sons, 1943.

[54] Y. Xia, S. C. Douglas and D. P. Mandic, "Adaptive frequency estimation in smart grid applications: Exploiting noncircularity and widely linear adaptive estimators," *IEEE Signal Processing Magazine*, vol. 29, no. 5, pp. 44–54, September 2012.

[55] S. Kanna, S. P . Talebi, and D. P. Mandic, "Diffusion widely linear adaptive estimation of system frequency in distributed power grids," *IEEE International Energy Conference*, pp. 772–778, 2014.

[56] Y. Xia and D. P. Mandic, "A widely linear least mean phase algorithm for adaptive frequency estimation of unbalanced power systems," *International Journal of Electrical Power & Energy Systems*, vol. 54, pp. 367–375, 2014.

[57] Y. Xia, S. C. Douglas, and D. P. Mandic, "Widely linear adaptive frequency estimation in three-phase power systems under unbalanced voltage sag conditions," *In Proceedings of International Joint Conference on Neural Networks*, pp. 1700–1705, July 2011.

[58] D. H. Dini, Y. Xia, S. C. Douglas, and D. P. Mandic, "Widely linear state space models for frequency estimation in unbalanced three-phase systems," *IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 9–12, June 2012.

[59] W. R. Hamilton, *Elements of quaternions.* Longmans & Co., 1866.

[60] J. B. Kuipers, *Quaternions and rotation sequences: A primer with applications to orbits, aerospace and virtual reality.* Princeton University Press, August 2002.

[61] K. W. Spring, "Euler parameters and the use of quaternion algebra in the manipulation of finite rotations: A review," *Mechanism and Machine Theory*, vol. 21, no. 5, pp. 365–373, 1986.

[62] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 1, pp. 12–28, January 2007.

[63] X. Yun, M. Lizarraga, E. R. Bachmann, and R. B. McGhee, "An improved quaternion-based Kalman filter for real-time tracking of rigid body orientation," *In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2, pp. 1074–1079, October 2003.

[64] J. Lee and S. Y. Shin, "General construction of time-domain filters for orientation data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 2, pp. 119–128, April 2002.

[65] S. J. Sangwine, “Fourier transforms of colour images using quaternion, or hypercomplex, numbers,” *IEEE Electronic Letters*, vol. 32, no. 21, pp. 1979–1980, October 1996.

[66] S. C. Pei and C. M. Cheng, “Color image processing by using binary quaternion-moment-preserving thresholding technique,” *IEEE Transactions on Image Processing*, vol. 8, no. 5, pp. 614–628, May 1999.

[67] T. A. Ell and S. J. Sangwine, “Hypercomplex Fourier transforms of color images,” *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 22–35, January 2007.

[68] S. Miron, N. Le Bihan, and J. I. Mars, “Quaternion-MUSIC for vector-sensor array processing,” *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1218–1229, April 2006.

[69] D. P. Mandic, C. Jahanchahi, and C. C. Took, “A quaternion gradient operator and its applications,” *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 47–50, January 2011.

[70] C. C. Took and D. P. Mandic, “A quaternion widely linear adaptive filter,” *IEEE Transactions on Signal Processing*, vol. 58, no. 8, pp. 4427–4431, August 2010.

[71] C. C. Took and D. P. Mandic, “Augmented second-order statistics of quaternion random signals,” *Signal Processing*, vol. 91, no. 2, pp. 214–224, 2011.

[72] S. Said, N. Le Bihan, and S. J. Sangwine, “Fast complexified quaternion Fourier transform,” *IEEE Transactions on Signal Processing*, vol. 56, no. 4, pp. 1522–1531, April 2008.

[73] N. Le Bihan, S. J. Sangwine, and T. A. Ell, “Instantaneous frequency and amplitude of orthocomplex modulated signals based on quaternion fourier transform,” *Signal Processing*, vol. 94, pp. 308–318, 2014.

[74] F. A. Tobar and D. P. Mandic, “Quaternion reproducing kernel Hilbert spaces: Existence and uniqueness conditions,” *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5736–5749.

[75] C. C. Took, G. Strbac, K. Aihara, and D. P. Mandic, “Quaternion valued short-term joint forecasting of three-dimensional wind and atmospheric parameters,” *Renewable Energy*, vol. 36, no. 6, pp. 1754–1760, June 2011.

[76] D. H. Dini, C. Jahanchahi, and D. P. Mandic, “Kalman filtering for widely linear complex and quaternion valued bearings only tracking,” *IET Signal Processing*, vol. 6, no. 5, pp. 435–445, July 2012.

[77] J. Via, D. Ramirez, and I. Santamaria, “Properness and widely linear processing of quaternion random vectors,” *IEEE Transactions on Information Theory*, vol. 56, no. 7, pp. 3502–3515, July 2010.

[78] J. Via, D. Ramirez, I. Santamaria, and L. Vielva, "Widely and semi-widely linear processing of quaternion vectors," *In Proceeidngs of IEEE International Confrence on Acoustics, Speech, and Signal Processing*, pp. 3946–3949, March 2010.

[79] C. Jahanchahi, C. C. Took, and D. P. Mandic, "On $\mathbb{HR}$-calculus, quaternion valued stochastic gradient, and adaptive three dimensional wind forecasting," *International Joint Conference on Neural Networks*, pp. 1–5, July 2010.

[80] E. Kreyszic, H. Kreyszic, and E. J. Norminton, *Advanced engineering mathmatics*, 10th ed. John Wiley & Sons, 2010.

[81] K. Kreutz–Delgado, *The complex gradient operator and the $\mathbb{CR}$-calculus.* University of California, San Diego, Technical Report, 2006.

[82] B. icinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Tansactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, August 1995.

[83] P. O. Amblard, M. Gaeta, and J. L. Lacoume, "Statistics for complex variables and signals — part I: Variables," *Signal Processing*, vol. 53, no. 1, pp. 1–13, 1996.

[84] T. A. Ell and S. J. Sangwine, "Quaternion involutions and anti-involutions," *Computers & Mathematics with Applications*, vol. 53, no. 1, pp. 137–143, 2007.

[85] C. Jahanchahi, "Quaternion valued adaptive signal processing," Ph.D. dissertation, Imperial College London, 2013.

[86] K. Shoemake, "Animating rotation with quaternion curves," *SIGGRAPH Computer Graphics*, vol. 19, no. 3, pp. 245–254, July 1985.

[87] C. Jahanchahi and D. P. Mandic, "A class of quaternion *Kalman* filters," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 3, pp. 533–544, March 2014.

[88] C. C. Took and D. P. Mandic, "The quaternion LMS algorithm for adaptive filtering of hypercomplex processes," *IEEE Transactions on Signal Processing*, vol. 57, no. 4, pp. 1316–1327, April 2009.

[89] D. Choukroun, I. Bar-Itzhack, and Y. Oshman, "Novel quaternion Kalman filter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 174–190, January 2006.

[90] J. Marins, X. Yun, E. Bachmann, R. McGhee, and M. Zyda, "An extended Kalman filter for quaternion-based orientation estimation using MARG sensors," *In Proceedings of International Conference on Intelligent Robots and Systems*, vol. 4, pp. 2003–2011, 2001.

[91] X. Yun and E. R. Bachmann, “Design, implementation, and experimental results of a quaternion-based Kalman filter for human body motion tracking,” *IEEE Transactions on Robotics*, vol. 22, no. 6, pp. 1216–1227, December 2006.

[92] A. M. Sabatini, “Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing,” *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 7, pp. 1346–1356, July 2006.

[93] E. Ollila, “On the circularity of a complex random variable,” *IEEE Signal Processing Letters*, vol. 15, pp. 841–844, 2008.

[94] S. Kanna, S. C. Douglas, and D. P. Mandic, “A real time tracker of complex circularity,” *In Proceedings of IEEE Sensor Array and Multichannel Signal Processing Workshop*, pp. 129–132, June 2014.

[95] P. O. Amblard, M. Gaeta, and J. L. Lacoume, “Statistics for complex variables and signals — part II: Signals,” *Signal Processing*, vol. 53, no. 1, pp. 15–25, 1996.

[96] N. N. Vakhania, “Random vectors with values in the quaternion hilbert spaces,” *Theory of Probability and Its Applications*, vol. 43, no. 1, pp. 99–115, 1999.

[97] P. O. Amblard and N. L. Bihan, “On quaternion valued random variables,” *In Proceedings of International Conference on Mathematics in Signal processing*, pp. 23–26, 2006.

[98] J. Via, D. P. Palomar, and L. Vielva, “Generalized likelihood ratios for testing the properness of quaternion Gaussian vectors,” *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1356–1370, April 2011.

[99] S. Javidi, C. C. Took, and D. P. Mandic, “Fast independent component analysis algorithm for quaternion valued signals,” *IEEE Transactions on Neural Networks*, vol. 22, no. 12, pp. 1967–1978, December 2011.

[100] C. C. Took, C. Jahanchahi, and D. P. Mandic, “A unifying framework for the analysis of quaternion valued adaptive filters,” *In Proceedings of Forty Fifth Asilomar Conference on Signals, Systems, and Computers*, pp. 1771–1774, November 2011.

[101] S. M. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, October 1998.

[102] J. Via, D. P. Palomar, L. Vielva, and I. Santamaria, “Quaternion ICA from second-order statistics,” *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1586–1600, April 2011.

[103] P. J. Schreier, L. L. Scharf, and A. Hanssen, "A generalized likelihood ratio test for impropriety of complex signals," *IEEE Signal Processing Letters*, vol. 13, pp. 433–436, July 2006.

[104] A. Tarighat and A. H. Sayed, "Least mean-phase adaptive filters with application to communications systems," *IEEE Signal Processing Letters*, vol. 11, no. 2, pp. 220–223, February 2004.

[105] S. C. Douglas and D. P. Mandic, "The least-mean-magnitude-phase algorithm with applications to communications systems," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 4152–4155, May 2011.

[106] A. Baggini, *Handbook of power quality.* John Wiley & Sons, 2008.

[107] R. K. Varma, R. M. Mathur, G. J. Rogers, and P. Kundur, "Modeling effects of system frequency variation in long-term stability studies," *IEEE Transactions on Power Systems*, vol. 11, no. 2, pp. 827–832, May 1996.

[108] M. H. J. Bollen, "Voltage sags in three-phase systems," *IEEE Power Engineering Review*, vol. 21, no. 9, pp. 8–15, September 2001.

[109] S. Chung, "A phase tracking system for three phase utility interface inverters," *IEEE Transactions on Power Electronics*, vol. 15, no. 3, pp. 431–438, May 2000.

[110] H. Karimi, M. Karimi-Ghartemani, and M. R. Iravani, "Estimation of frequency and its rate of change for applications in power systems," *IEEE Transactions on Power Delivery*, vol. 19, no. 2, pp. 472–480, April 2004.

[111] V. V. Terzija, "Improved recursive Newton-type algorithm for frequency and spectra estimation in power systems," *IEEE Transaction on Instrumentation and Measurement*, vol. 52, no. 5, pp. 1654–1659, October 2003.

[112] V. V. Terzija, B. N. Djuric, and B. D. Kovacevic, "Voltage phasor and local system frequency estimation using Newton type algorithm," *IEEE Transactions on Power Delivery*, vol. 9, no. 3, pp. 1368–1374, July 1994.

[113] A. Abdollahi and F. Matinfar, "Frequency estimation: A least-squares new approach," *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 790–798, April 2011.

[114] A. K. Pradhan, A. Routray, and A. Basak, "Power system frequency estimation using least mean square technique," *IEEE Transactions on Power Delivery*, vol. 20, no. 3, pp. 1812–1816, July 2005.

[115] T. Lobos and J. Rezmer, "Real-time determination of power system frequency," *IEEE Transactions on Instrumentation and Measurement*, vol. 46, no. 4, pp. 877–881, August 1997.

[116] M. Mojiri, D. Yazdani, and A. Bakhshai, “Robust adaptive frequency estimation of three-phase power systems,” *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 7, pp. 1793–1802, July 2010.

[117] P. K. Dash, R. K. Jena, G. Panda, and A. Routray, “An extended complex Kalman filter for frequency measurement of distorted signals,” *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 4, pp. 746–753, August 2000.

[118] A. Routray, A. K. Pradhan, K. P. Rao, “A novel Kalman filter for frequency estimation of distorted signals in power systems,” *IEEE Transactions on Instrumentation and Measurement*, vol. 51, no. 3, pp. 469–479, June 2002.

[119] P. K. Dash, A. K. Pradhan, and G. Panda, “Frequency estimation of distorted power system signals using extended complex Kalman filter,” *IEEE Transactions on Power Delivery*, vol. 14, no. 3, pp. 761–766, July 1999.

[120] H. S. Song and K. Nam, “Instantaneous phase-angle estimation algorithm under unbalanced voltage-sag conditions,” *IEE Proceedings - Generation,Transmission and Distribution*, vol. 147, no. 6, pp. 409–415, November 2000.

[121] Y. Xia and D. P. Mandic, “Widely linear adaptive frequency estimation of unbalanced three-phase power systems,” *IEEE Transactions on Instrumentation and Measurement*, vol. 61, no. 1, pp. 74–83, Jan 2012.

[122] D. H. Dini and D. P. Mandic, “Widely linear modeling for frequency estimation in unbalanced three-phase power systems,” *IEEE Transactions on Instrumentation and Measurement*, vol. 62, no. 2, pp. 353–363, February 2013.

[123] J. Arrillaga and N. R. Watson, *Power system harmonics*, 2nd ed. Wiley, 2003.

[124] R. Olfati-Saber, “Distributed tracking for mobile sensor networks with information-driven mobility,” *In Proceeding of The American Control Conference*, pp. 4604–4612, July 2007.

[125] R. Olfati-Saber and R. M. Murray, “Consensus problems in networks of agents with switching topology and time-delays,” *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, September 2004.

[126] E. Franco, R. Olfati-Saber, T. Parisini, and M. M. Polycarpou, “Distributed fault diagnosis using sensor networks and consensus-based filters,” *In Proceedings of IEEE Conference on Decision & Control*, pp. 386–391, December 2006.

[127] R. Olfati-Saber, “Flocking for multi-agent dynamic systems: Algorithms and theory,” *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, March 2006.

[128] R. Olfati-Saber and P. Jalalkamali, "Coupled distributed estimation and control for mobile sensor networks," *IEEE Transactions on Automatic Control*, vol. 57, no. 10, pp. 2609–2614, Oct 2012.

[129] F. Cattivelli and A. H. Sayed, "Distributed nonlinear Kalman filtering with applications to wireless localization," *In Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3522–3525, March 2010.

[130] F. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, September 2010.

[131] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Foundations and Trends in Machine Learning*, vol. 7, no. 4-5, pp. 311–801, 2014.

[132] F. S. Cattivelli and A. H. Sayed, "Diffusion LMS strategies for distributed estimation," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1035–1048, March 2010.

[133] C. Jianshu and A. H. Sayed, "Diffusion adaptation strategies for distributed optimization and learning over networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4289–4305, August 2012.

[134] P. Braca, P. Willett, K. Le Page, S. Marano, and V. Matta, "Bayesian tracking in underwater wireless sensor networks with port-starboard ambiguity," vol. 62, no. 7, pp. 1864–1878, April 2014.

[135] T Kailath, A. H. Sayed, and B. Hassibi, *Linear estimation.* Prentice Hall, 2000.

[136] X. Shen, H. Zhang, Y. Xu, and S. Meng, "Observation of alpha-stable noise in the laser gyroscope data," *IEEE Sensors Journal*, vol. 16, no. 7, pp. 1998–2003, April 2016.

[137] M. Shao and C. L. Nikias, "Signal processing with fractional lower order moments: Stable processes and their applications," *In Proceedings of The IEEE*, vol. 81, no. 7, pp. 986–1010, 1993.

[138] G. Samorodnitsky and M. S. Taqqu, *Stable non-Gaussian random processes: Stochastic models with infinite variance.* *Chapman & Hall*, 1994.

[139] V. V. Uchaikin and V. M. Zolotarev, *Chance and stability: Stable distributions and their applications.* *Walter de Gruyter*, 1999.

[140] J. S. Bodenschatz, "Symmetric alpha-stable filter theory," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, pp. 1945–1948, 1997.

[141] J. P. Nolan and D Ojeda-Revah, "Linear and nonlinear regression with stable errors," *Journal of Econometrics*, vol. 172, no. 2, pp. 186–194, February 2013.

[142] G. Aydin, O. Arika, and A. E. Cetin, "Robust adaptive filtering algorithms for α-stable random processes," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*.

[143] O. Arikan, A. E. Cetin, and E. Erzin, "Adaptive filtering for non-gaussian stable processes," *IEEE Signal Processing Letters*, vol. 1, no. 11, pp. 163–165, November 1994.

[144] E. Masry, "Alpha-stable signals and adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 48, no. 11, pp. 3011–3016, November 2000.

[145] R. Sharma, W. A. Sethares, and J. A. Bucklew, "Asymptotic analysis of stochastic gradient-based adaptive filtering algorithms with general cost functions," *IEEE Transactions on Signal Processing*, vol. 44, p. 2186–2194, 1996.

[146] I. A. Koutrouvelis, "An iterative procedure for the estimation of the parameters of stable laws," *Communications in Statistics: Simulation and Computation*, vol. IO, no. 1, pp. 17–28, 1981.

[147] B. Ristic, S. Arulampalam, and N. Gordon, *Beyond the Kalman filter: Particle filters for tracking applications*. Artech House, 2004.

[148] D. H. Dini, P. M. Djuric, and D. P. Mandic, "The augmented complex particle filter," *IEEE Transactions on Signal Processing*, vol. 61, no. 17, pp. 4341–4346, September 2013.

[149] A. Papoulis, *Probability, random variables, and stochastic processes*, 3rd ed. McGraw-Hill, 1991.

Appendix A

Proof of Equation 3.13

Given a quaternion-valued zero-mean Gaussian random variable $q = q_r + iq_i + jq_j + kq_k$; then, $R_{qq^{i*}}$ can be expressed as

$$\begin{aligned} R_{qq^{i*}} = E[qq^{i*}] =& E[q_r q_r] + E[q_i q_i] - E[q_j q_j] - E[q_k q_k] \\ & + 2j\left(E[q_r q_j] - E[q_i q_k]\right) + 2k\left(E[q_r q_k] + E[q_j q_i]\right) \\ =& C_{q_r} + C_{q_i} - C_{q_j} - Cq_k \\ & + 2j\left(R_{q_r q_j} - R_{q_i q_k}\right) + 2k\left(R_{q_r q_k} + R_{q_j q_i}\right). \end{aligned} \tag{A.1}$$

Moreover, from the expression in (A.1) we have

$$\begin{aligned} |R_{qq^{i*}}|^2 =& C_{q_r}^2 + C_{q_i}^2 + C_{q_j}^2 + C_{q_k}^2 + 2C_{q_r}C_{q_i} + 2C_{q_j}C_{q_k} \\ & - 2C_{q_r}C_{q_j} - 2C_{q_r}C_{q_k} - 2C_{q_k}C_{q_j} - 2C_{q_i}C_{q_k} \\ & + 4R_{q_r q_j}^2 + 4R_{q_i q_k}^2 - 8R_{q_r q_j}R_{q_i q_k} \\ & + 4R_{q_r q_k}^2 + 4R_{q_i q_j}^2 + 8R_{q_r q_k}R_{q_i q_k}. \end{aligned} \tag{A.2}$$

Furthermore, the expression for $|R_{qq^{j*}}|^2$ and $|R_{qq^{k*}}|^2$ can be formulated in an analogous manner as

$$\begin{aligned} |R_{qq^{j*}}|^2 =& C_{q_r}^2 + C_{q_i}^2 + C_{q_j}^2 + C_{q_k}^2 - 2C_{q_r}C_{q_i} + 2C_{q_r}C_{q_j} \\ & - 2C_{q_r}C_{q_k} - 2C_{q_i}C_{q_j} + 2C_{q_i}C_{q_k} - 2C_{q_j}C_{q_k} \\ & + 4R_{q_r q_i}^2 + 4R_{q_j q_k}^2 + 8R_{q_r q_i}R_{q_j q_k} \\ & + 4R_{q_r q_k}^2 + 4R_{q_i q_j}^2 - 8R_{q_r q_k}R_{q_i q_j} \end{aligned} \tag{A.3}$$

$$\begin{aligned} |R_{qq^{k*}}|^2 =& C_{q_r}^2 + C_{q_i}^2 + C_{q_j}^2 + C_{q_k}^2 - 2C_{q_r}C_{q_i} - 2C_{q_r}C_{q_j} \\ & + 2C_{q_r}C_{q_k} + 2C_{q_i}C_{q_j} - 2C_{q_i}C_{q_k} - 2C_{q_j}C_{q_k} \\ & + 4R_{q_r q_i}^2 + 4R_{q_k q_j}^2 - 8R_{q_r q_i}R_{q_k q_j} \\ & + 4R_{q_r q_k}^2 + 4R_{q_r q_j}^2 + 8R_{q_i q_k}R_{q_r q_j}. \end{aligned} \tag{A.4}$$

In addition, taking into account that $C_q = E[qq^*] = C_{q_r} + C_{q_i} + C_{q_j} + C_{q_k}$, we have

$$\begin{aligned} |C_q|^2 = &\left(C_{q_r}^2 + C_{q_i}^2 + C_{q_j}^2 + C_{q_k}^2\right) \\ &+ 2\left(C_{q_r}C_{q_i} + C_{q_r}C_{q_j} + C_{q_r}C_{q_k} + C_{q_i}C_{q_j} + C_{q_j}C_{q_k}\right). \end{aligned} \tag{A.5}$$

On the other hand $E\left[|qq^*|^2\right] = E\left[\left|q_r^2 + q_i^2 + q_j^2 + q_k^2\right|^2\right]$, which after some tedious mathematical manipulations can be simplified to give

$$\begin{aligned} E\left[|qq^*|^2\right] =&3\left(C_{q_r}^2 + C_{q_i}^2 + C_{q_j}^2 + C_{q_k}^2\right) \\ &+ 4\left(R_{q_rq_i}^2 + R_{q_rq_j}^2 + R_{q_rq_k}^2 + R_{q_iq_j}^2 + R_{q_jq_k}^2\right) \\ &- 2\left(C_{q_r}C_{q_i} + C_{q_r}C_{q_j} + C_{q_r}C_{q_k} + C_{q_i}C_{q_j} + C_{q_j}C_{q_k}\right). \end{aligned} \tag{A.6}$$

Now from comparing the expressions in (A.2), (A.3), (A.4), and (A.5) with the expression in(A.6) it becomes clear that

$$E\left[|qq^*|^2\right] = \frac{3|C_q|^2 + |R_{qq^i}|^2 + |R_{qq^j}|^2 + |R_{qq^k}|^2}{2} = \frac{C_{q_n}^2}{2}\left(3 + |\rho_i|^2 + |\rho_j|^2 + |\rho_k|^2\right).$$

Appendix B

Implementation of the Quaternion Hilbert Transform

A general ellipse in the i-j plane is given by

$$q_n = iA\cos\left(2\pi f \Delta T n + \phi_1\right) + jB\sin\left(2\pi f \Delta T n + \phi_2\right)$$

where $\{A, B\} \in \mathbb{R}$ are arbitrary amplitudes and $\{\phi_1, \phi_2\} \in [0, 2\pi)$ are arbitrary phases. The Hilbert transform of q_n with respect to the imaginary unit k is given by

$$\mathcal{H}_k(q_n) = iAk\sin\left(2\pi f \Delta T n + \varphi_1\right) - jBk\cos\left(2\pi f \Delta T n + \varphi_2\right).$$

Using simple mathematical manipulations the following signal is constructed

$$S_n = q_n + H_k(q_n) = \left(iAe^{k\varphi_1} - jBe^{k\varphi_2}\right) e^{k2\pi f \Delta T n}.$$

It can be shown that

$$\frac{S_{n+1}}{S_n} = e^{k2\pi f \Delta T}. \tag{B.1}$$

The proof for the case of Hilbert transforms with respect to i and j-axis follows similarly. Notice that it was assumed that the amplitude of S_n does not change significantly over one time interval; thus, in (B.1) S_n and S_{n+1} can be normalized. For the purpose of simulations, the Hilbert transform was implemented using a finite impulse response filter with 30 taps.

Appendix C

Deterministic Relation Between Voltage Phasors and the State Space Elements of the QFE

Without loss of generality phase "a" is selected as the reference for the phasor angles, i.e. it was assumed that $\phi_{a,n} = 0$, which yields

$$\begin{aligned} \Lambda_{I,n} &= iV_{a,n} + jV_{b,n}\cos(\phi_{b,n} + \frac{4\pi}{3}) + kV_{c,n}\cos(\phi_{c,n} + \frac{4\pi}{3}) \\ \Lambda_{Q,n} &= jV_{b,n}\sin(\phi_{b,n} + \frac{2\pi}{3}) + kV_{c,n}\sin(\phi_{c,n} + \frac{4\pi}{3}). \end{aligned} \tag{C.1}$$

From the expression in (C.1), notice that multiplying q_n by $v_{a,n}$ gives

$$q_n v_{a,n} = \frac{1}{2}\Lambda_{I,n}V_{a,n} + \frac{1}{2}\Lambda_{I,n}V_{a,n}\cos(4\pi f\Delta Tn) - \frac{1}{2}\Lambda_{Q,n}V_{a,n}\sin(4\pi f\Delta Tn). \tag{C.2}$$

Thus, $\Lambda_{I,n}V_{a,n}$ can be calculated by passing $q_n v_{a,n}$, as expressed in (C.2), through a low pass filter (LPF) which can be expressed as

$$2q_n v_{a,n} \rightarrow \boxed{\text{LPF}} \rightarrow \kappa_n = \Lambda_{I,n}V_{a,n}$$

It was assumed that $\phi_{a,n} = 0$; therefore, it follows that

$$\Lambda_{I,n} = \frac{\kappa_n}{\sqrt{-i\Im_i(\kappa_n)}}. \tag{C.3}$$

Taking into account the expression in (6.9) and applying tedious mathematical manipulations it can be shown that

$$q_n^+ = (\frac{\Lambda_{I,n}}{2} - \frac{\Lambda_{Q,n}}{2\zeta}\Big)e^{(\zeta 2\pi f\Delta Tn)} \quad \text{and} \quad q_n^- = \Big(\frac{\Lambda_{I,n}}{2} + \frac{\Lambda_{Q,n}}{2\zeta}\Big)e^{-(\zeta 2\pi f\Delta Tn)}. \tag{C.4}$$

From the expressions in (C.4) we have

$$|q_n^+|^2 - |q_n^-|^2 = |\Lambda_{I,n}||\Lambda_{Q,n}|\Re(\lambda_{I,n}\lambda_{Q,n}\zeta)$$

where $\lambda_{I,n} = \Lambda_{I,n}/|\Lambda_{I,n}|$ and $\lambda_{Q,n} = \Lambda_{Q,n}/|\Lambda_{Q,n}|$. Considering that ζ in normal to $\lambda_{I,n}$ and $\lambda_{Q,n}$ yields

$$\lambda_{I,n}\lambda_{Q,n}\zeta = \underbrace{(\lambda_{I,n} \cdot \lambda_{Q,n})\zeta}_{\text{pure quaternion}} - \underbrace{(\lambda_{I,n} \times \lambda_{Q,n})\zeta}_{\sin(\phi_{I,Q})}$$

where $\phi_{I,Q}$ is the angle between $\Lambda_{I,n}$ and $\Lambda_{Q,n}$; therefore,

$$|q_n^+|^2 - |q_n^-|^2 = -|\Lambda_{I,n}||\Lambda_{Q,n}|\sin(\phi_{I,Q}). \tag{C.5}$$

Furthermore, using analytical geometry it can be show that

$$\begin{aligned} \Lambda_{I,n} \times \Lambda_{Q,n} =& - |\Lambda_{I,n}||\Lambda_{Q,n}|\sin(\phi_{I,Q})\zeta \\ =& iV_{b,n}V_{c,n}\sin(\frac{2\pi}{3} + \phi_{c,n} - \phi_{b,n}) \\ & - jV_{a,n}V_{c,n}\sin(\frac{4\pi}{3} + \phi_{c,n}) \\ & + kV_{a,n}V_{b,n}\sin(\frac{2\pi}{3} - \phi_{b,n}) \end{aligned} \tag{C.6}$$

and therefore replacing (C.5) into (C.6) yields

$$\begin{aligned} \Lambda_{Q,n} =& i\Im_k\left(\left(|q_n^+|^2 - |q_n^-|^2\right)\zeta\right)\left(\frac{1}{\sqrt{-i\Im_i(\kappa_n)}}\right) \\ & - i\Im_j\left(\left(|q_n^+|^2 - |q_n^-|^2\right)\zeta\right)\left(\frac{1}{\sqrt{-i\Im_i(\kappa_n)}}\right) \end{aligned} \tag{C.7}$$

where ζ can be obtained through $\Im(\varphi_n)/|\Im(\varphi_n)|$.

Appendix D

The Relation Between the PDF and Circularity

Although the approach based on the AQCF presented in Section 8.3.4 is general, as the AQCF is guaranteed to exist for all quaternion-valued random variables and the AQCF of SαS random variables confined to the format given in Section 8.4, for the sake completeness, akin to approaches to symmetricity in $\mathbb{R}$ [149], the relation between the pdf of quaternion-valued random variables and circularity is investigated in the sequel.

Consider a quaternion-valued random vector, $\boldsymbol{Q}$, with i.i.d components, the pdf of which exits and can be expressed as

$$P_{\boldsymbol{Q}^a}(\mathbf{q}^a) = P_{\boldsymbol{Q}_r}(\mathbf{q}_r)P_{\boldsymbol{Q}_i}(\mathbf{q}_i)P_{\boldsymbol{Q}_j}(\mathbf{q}_j)P_{\boldsymbol{Q}_k}(\mathbf{q}_k) = f(l)$$

where l is given by

$$l = \sqrt{\mathbf{q}_r^T\mathbf{G}_r\mathbf{q}_r + \mathbf{q}_i^T\mathbf{G}_i\mathbf{q}_i + \mathbf{q}_j^T\mathbf{G}_j\mathbf{q}_j + \mathbf{q}_k^T\mathbf{G}_k\mathbf{q}_k} \tag{D.1}$$

and f$(\cdot) : \mathbb{R} \to \mathbb{R}$. Making the assumption that $f(\cdot)$ and $P_{\boldsymbol{Q}_r}$ are differentiable yields

$$\frac{\partial f(l)}{\partial \mathbf{q}_r} = \frac{\partial f(l)}{\partial l}\frac{\partial l}{\partial \mathbf{q}_r} = \frac{\partial P_{\boldsymbol{Q}_r}(\mathbf{q}_r)}{\partial \mathbf{q}_r}P_{\boldsymbol{Q}_i}(\mathbf{q}_i)P_{\boldsymbol{Q}_j}(\mathbf{q}_j)P_{\boldsymbol{Q}_k}(\mathbf{q}_k). \tag{D.2}$$

Through straightforward mathematical manipulations the expression in (D.2) gives

$$\frac{\partial f(l)}{\partial l}\frac{\mathbf{G}_r\mathbf{q}_r}{l} = \frac{\partial P_{\boldsymbol{Q}_r}(\mathbf{q}_r)}{\partial \mathbf{q}_r}P_{\boldsymbol{Q}_i}(\mathbf{q}_i)P_{\boldsymbol{Q}_j}(\mathbf{q}_j)P_{\boldsymbol{Q}_k}(\mathbf{q}_k). \tag{D.3}$$

Now, making the assumption that $\mathbf{G}$ is positive definite and dividing both sides of the expression in (D.3) by $f(l)$ yields

$$\frac{\partial f(l)}{\partial l}\frac{1}{lf(l)} = \frac{\mathbf{q}_r^T}{\mathbf{q}_r^T\mathbf{G}_r\mathbf{q}_r}\frac{\partial P_{\boldsymbol{Q}_r}(\mathbf{q}_r)}{\partial \mathbf{q}_r}. \tag{D.4}$$

Note that the right hand side of the equation in (D.4) is only dependent on $\mathbf{q}_r$ whereas its left hand side is dependent on l, a function of $\{\mathbf{q}_r, \mathbf{q}_i, \mathbf{q}_j, \mathbf{q}_k\}$; therefore, the right hand side of the equation in (D.4) must be a constant resulting in

$$\frac{\partial f(l)}{\partial l}\frac{1}{f(l)} = \frac{\partial \ln(f(l))}{\partial l} = \underbrace{\left(\frac{\mathbf{q}_r^T}{\mathbf{q}_r^T\mathbf{G}_r\mathbf{q}_r}\frac{\partial P_{\mathbf{Q}_r}(\mathbf{q}_r)}{\partial \mathbf{q}_r}\right)}_{\gamma} l \tag{D.5}$$

where solving the differential equation in (D.5) gives

$$f(d) = ae^{\frac{\gamma}{2}l^2} \tag{D.6}$$

with $a \in \mathbb{R}^+$ selected such that

$$\int_{\mathcal{D}^a_{\mathbf{Q}}} P_{\boldsymbol{Q}^a}(\mathbf{q}^a)\mathrm{d}\mathbf{q}^a = 1 \tag{D.7}$$

where $\mathcal{D}_{\boldsymbol{Q}^a}$ is the entire domain of $\boldsymbol{Q}$, in order to grantee that the expression in (D.6) is a pdf. Substituting (D.1) into (D.6), using the transform in (2.8), and selecting a such that the condition in (D.7) is satisfied yields

$$f(d) = \frac{e^{\frac{-1}{2}\mathbf{q}^{aH}\mathcal{G}\mathbf{q}^a}}{(\pi^2/4)^N\sqrt{\det(\mathcal{G})}} \tag{D.8}$$

where $\mathcal{G}$ is a $4N \times 4N$ block diagonal matrix given by

$$\mathcal{G} = \begin{bmatrix} \mathbf{G} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G} \end{bmatrix} \quad \text{where} \quad \mathbf{G} = \frac{1}{\gamma}\left(\mathbf{G}_r^{-1} + \mathbf{G}_i^{-1} + \mathbf{G}_j^{-1} + \mathbf{G}_k^{-1}\right)^{-1}.$$

Note that the distribution in (D.8) is a Gaussian distribution with augmented covariance matrix $\mathcal{G}^{-1}$. In addition, if $\mathbf{G}_r = \mathbf{G}_i = \mathbf{G}_j = \mathbf{G}_k = \vartheta\mathbf{I}$, where $\vartheta \in \mathbb{R}^+$, in the expression in (D.8), $f(d)$ will simplify into a circular Gaussian distribution. Therefore, if the components of $\boldsymbol{Q}$ are i.i.d, the resulting distribution is circular if and only if the real-valued components of $\boldsymbol{Q}$ are Gaussian random vectors with the same diagonal covariance matrices.