The Quaternion Least Mean Magnitude Phase Adaptive Filtering Algorithm

Min Xiang[1], Scott C. Douglas[2], and Danilo P. Mandic[1]
[1]Department of Electrical and Electronic Engineering, Imperial College London, SW7 2BT, U.K.
[2]Department of Electrical Engineering, Southern Methodist University, Dallas, Texas 75275, USA
Emails: {m.xiang13, d.mandic}@ic.ac.uk, douglas@lyle.smu.edu

***Abstract*—Quaternion-valued adaptive filters based on the mean square error (MSE) criterion have been extensively studied in recent years. However, the MSE cost function has only one degree of freedom, and to circumvent this problem, we propose another criterion which enables separate control of the magnitude and phase. Next, a quaternion least mean magnitude phase (QLMMP) filtering algorithm is introduced and is shown to provide a unified form for the quaternion least mean square (QLMS) algorithm, a magnitude-only filtering algorithm, and a phase filtering algorithm. The convergence analysis of the normalised version of the QLMMP employs a connection to the normalised QLMS with a quaternion-valued stepsize. The proposed algorithms are validated over case studies of the identification in three-phase power systems.**

***Index Terms*—Quaternions, adaptive filtering, magnitude, phase, convergence analysis.**

I. Introduction

Quaternions have traditionally been used in aerospace engineering and computer graphics to model three-dimensional rotation and orientation, as their algebra avoids numerical problems associated with vector algebras [1]. The recently introduced HR calculus [2], [3] and augmented quaternion statistics [4], [5] have triggered a resurgence of research on quaternion-valued signal processing, owing to a compact model of mutual interaction between the data channels provided by quaternions, and the inherent physically meaningful interpretation for a number of three-dimensional and four-dimensional problems [6]. Quaternions have subsequently found new applications in areas including communications, motion tracking, and biomedical signal processing [7]–[10].

The minimisation of mean square error (MSE) has been widely used as the basis for the development of adaptive filtering algorithms. The least mean square (LMS) algorithm is the most popular example of adaptive MSE minimisation, and has been extended from the real domain to complex and quaternion domains in order to cater for multidimensional signals [11]–[14]. However, the LMS is known not to perform most efficiently for complex and quaternion signals with different dynamics in their magnitudes and phases, such as those caused by Doppler effects [15]. To improve the filtering performance, several criteria have been proposed to separately extract the magnitude and/or phase information from complex-valued signals, such as the least mean phase (LMP) algorithm for phase estimation [15], the constant modulus channel estimator for magnitude estimation [16], and the least mean magnitude phase (LMMP) algorithm for advanced estimation of magnitude and phase variations [17], [18].

Magnitude/phase coupling is inherent to quaternion-valued signals. Based on a definition of phase in the quaternion domain, a quaternion LMP (QLMP) algorithm has been introduced [19]. However, it requires complicated computations and the phase change in each iteration is restricted to be less than $\pi/2$. To this end, we here propose a quaternion LMMP (QLMMP) filtering algorithm which decomposes the instantaneous squared error cost into a composite of the squared magnitude error and a term representing the phase error between the desired signal and the estimate. In this way, a full control of magnitude and phase cost is achieved, together with the corresponding weight updates. We next provide convergence analysis for the normalised QLMMP algorithm by regarding the QLMMP as a special case of quaternion LMS (QLMS) with an adaptive quaternion-valued stepsize. This provides a very general algorithm. By setting equal stepsizes for the magnitude and phase update, the QLMMP is equivalent to the standard QLMS with a real-valued stepsize, while the vanishing phase (or magnitude) stepsize reduces the QLMMP to a magnitude-only (or phase) filtering algorithm. The phase filtering algorithm is shown to converge faster than the QLMP, and can be an alternative to the QLMS when combined with a simple magnitude adjustment. The performances of the proposed algorithms are validated by simulations in the context of frequency estimation of three-phase power systems.

II. Quaternion algebra

The quaternion domain $\mathbb{H}$ is a four-dimensional vector space over the real field spanned by the basis $\{1, \imath, \jmath, \kappa\}$ [20]. A quaternion variable $x \in \mathbb{H}$ consists of a scalar part $\Re\{\cdot\}$ and a vector part $\Im\{\cdot\}$ which comprises three imaginary components, so that $x = \Re[x] + \Im[x] = x_a + \imath x_b + \jmath x_c + \kappa x_d$, where $x_a, x_b, x_c, x_d \in \mathbb{R}$, and $\imath$, $\jmath$, κ are imaginary units with properties $\imath^2 = \jmath^2 = \kappa^2 = -1$, $\imath\jmath = -\jmath\imath = \kappa$, $\jmath\kappa = -\kappa\jmath = \imath$, $\kappa\imath = -\imath\kappa = \jmath$. The product of two quaternions x and y is

$$xy = \Re[x]\Re[y] - \Im[x]\Im[y] + \Re[x]\Im[y] + \Re[y]\Im[x] + \Im[x] \times \Im[y]$$

where the symbol $\times$ denotes vector product. The presence of the vector product causes the non-commutativity of the quaternion product, that is, $xy \neq yx$. The modulus of the quaternion is defined by $|x| = \left(x_a^2 + x_b^2 + x_c^2 + x_d^2\right)^{\frac{1}{2}}$. The quaternion conjugate operator $(\cdot)^*$ rotates the quaternion along all three imaginary axes, and is given by $x^* = \Re[x] - \Im[x] = x_a - \imath x_b - \jmath x_c - \kappa x_d$.

III. The quaternion LMMP (QLMMP)

For convenience, assume that the desired quaternion-valued signal at time n arises from the linear model, $d_n = \mathbf{w}_\mathrm{o}^T \mathbf{x}_n + \eta_n$, where $\mathbf{x}_n \in \mathbb{H}^{L\times 1}$ is the known input vector, $\mathbf{w}_\mathrm{o} \in \mathbb{H}^{L\times 1}$ is the unknown optimal weight vector of the system, and η_n is zero-mean white Gaussian noise (WGN) with variance σ_η^2. To estimate d_n as $y_n = \mathbf{w}_n^T \mathbf{x}_n$, where $\mathbf{w}_n \in \mathbb{H}^{L\times 1}$ is the weight vector, the QLMS algorithm, based on the minimisation of the instantaneous squared error cost $J_n = |d_n - y_n|^2$, has been proposed [13]. Using a stepsize μ_l and the gradient of J_n with respect to the conjugate of the weight vector, $\nabla_{\mathbf{w}_n^*} J_n = -(d_n - y_n)\,\mathbf{x}_n^*$, the weight update rule of the normalised QLMS is given by [3]

$$\mathbf{w}_{n+1} = \mathbf{w}_n - \frac{\mu_\mathrm{l}}{\|\mathbf{x}_n\|_2^2} \nabla_{\mathbf{w}_n^*} J_n = \mathbf{w}_n + \frac{\mu_\mathrm{l}}{\|\mathbf{x}_n\|_2^2} (d_n - y_n)\,\mathbf{x}_n^*. \quad (1)$$

Following the LMMP approach in [17], the squared error cost can be decomposed into

$$J_n = \underbrace{(|d_n| - |y_n|)^2}_{J_{\mathrm{m},n}} + \underbrace{2\,|d_n|\,|y_n|\,(1 - \cos\theta_n)}_{J_{\mathrm{p},n}} \quad (2)$$

where $J_{\mathrm{m},n}$ is the magnitude cost, $J_{\mathrm{p},n}$ the phase cost, $\theta_n \in [0, \pi]$ the absolute value of the angle between d_n and y_n, and $\cos\theta_n = \Re[d_n^* y_n]\,|d_n|^{-1}\,|y_n|^{-1}$. Similar to the complex LMMP algorithm [17], the minimisation of a linear weighted combination of $J_{\mathrm{m},n}$ and $J_{\mathrm{p},n}$, $J_{\mathrm{lmmp},n} = \alpha_\mathrm{m} J_{\mathrm{m},n} + \alpha_\mathrm{p} J_{\mathrm{p},n}$, where $\alpha_\mathrm{m}, \alpha_\mathrm{p} \in \mathbb{R}^+$ denote the weights, leads to a new algorithm referred to as the quaternion LMMP (QLMMP). Note that $J_{\mathrm{lmmp},n} \geq 0$, where the equality holds only when $d_n = y_n$. The gradients of $J_{\mathrm{lmmp},n}$ with respect to the conjugate of the weight vector are then calculated as [2], [3]

$$\nabla_{\mathbf{w}_n^*} J_{\mathrm{lmmp},n} = 0.5\left[\alpha_\mathrm{m}\left(1 - \frac{|d_n|}{|y_n|}\right) y_n + \alpha_\mathrm{p}\left(\frac{|d_n|}{|y_n|} y_n - d_n\right)\right]\mathbf{x}_n^*.$$

With the stepsize $\mu \in \mathbb{R}^+$, the weight update of the normalised QLMMP becomes

$$\begin{aligned} \mathbf{w}_{n+1} &= \mathbf{w}_n - \mu\,\|\mathbf{x}_n\|_2^{-2}\,\nabla_{\mathbf{w}_n^*} J_{\mathrm{lmmp},n} \\ &= \mathbf{w}_n + \tfrac{\mu_\mathrm{m}(|d_n| - |y_n|)}{\|\mathbf{x}_n\|_2^2 |y_n|} y_n \mathbf{x}_n^* + \tfrac{\mu_\mathrm{p}}{\|\mathbf{x}_n\|_2^2}\left(d_n - \tfrac{|d_n|}{|y_n|} y_n\right)\mathbf{x}_n^* \end{aligned} \quad (3)$$

where $\mu_\mathrm{m} = 0.5\mu\alpha_\mathrm{m}$ and $\mu_\mathrm{p} = 0.5\mu\alpha_\mathrm{p}$ control the respective convergence rates in the magnitude and phase, so that they are termed magnitude and phase stepsizes. By adjusting μ_m and μ_p, the QLMMP provides enhanced degrees of freedom in the minimisation of the cost function, and different convergence rates and accuracies in the magnitude and phase estimation. In contrast, the QLMS does not allow for such separate control of magnitude and phase.

From (3), given the new weight estimate $\mathbf{w}_{n+1}$, the improvement in the estimate for d_n is

$$\Delta y = \mathbf{w}_{n+1}^T \mathbf{x}_n - y_n = \underbrace{\mu_\mathrm{m}\left(\frac{|d_n|}{|y_n|} - 1\right) y_n}_{\Delta y_\mathrm{m}} + \underbrace{\mu_\mathrm{p}\left(d_n - \frac{|d_n|}{|y_n|} y_n\right)}_{\Delta y_\mathrm{p}}$$

where Δy_m aligns with y_n, and Δy_p rotates y_n towards d_n. Fig. 1 visualises the estimation improvements in an iteration of QLMS [3], QLMP [19], and the proposed QLMMP. On the plane where d_n and y_n lie, the estimation improvement in the QLMS, Δy_qlms, points to d_n, while the estimation improvement in the QLMP, Δy_qlmp, is perpendicular to y_n, indicating that the rotation angle of Δy_qlmp is always less than $\pi/2$. By weighting Δy_m and Δy_p via μ_m and μ_p, the estimation improvement in the QLMMP spans the shaded area where Δy_qlms and Δy_qlmp lie.

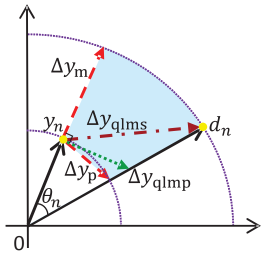


Fig. 1. Geometric interpretations of the QLMMP, QLMS, and QLMP. Although $\mathbb{H}$ is a four-dimensional vector space over $\mathbb{R}$, d_n, y_n and Δy all are on the same plane.

A. Convergence analysis

From (1) and (3), at time n, the QLMMP and QLMS are equivalent if the stepsize in (1) is quaternion-valued[1], that is,

$$\mu_\mathrm{l} = \mu_\mathrm{p} + \frac{(\mu_\mathrm{m} - \mu_\mathrm{p})\,(|d_n| - |y_n|)}{|d_n - y_n|^2\,|y_n|} y_n\,(d_n - y_n)^*. \quad (4)$$

This connection between the QLMMP and the QLMS with a quaternion-valued stepsize provides new insights into the convergence of the QLMMP, which is difficult to analyse straightforwardly, owing to the high nonlinearity of the cost function, $\mu_\mathrm{m} J_{\mathrm{m},n} + \mu_\mathrm{p} J_{\mathrm{p},n}$. For a set of μ_m and μ_p, if all values of μ_l calculated from (4) for all d_n and y_n fall into the convergence region of the normalised QLMS, the corresponding normalised QLMMP achieves a monotonic decrease in MSE. Therefore, the following analysis is based on the sufficient convergence condition of the normalised QLMS derived in the Appendix, that is, $2\Re[\mu_\mathrm{l}] - |\mu_\mathrm{l}|^2 > 0$.

From (4) we then obtain

$$\Re[\mu_\mathrm{l}] = \mu_\mathrm{p} + \frac{(\mu_\mathrm{m} - \mu_\mathrm{p})\,(|d_n| - |y_n|)}{|d_n - y_n|^2}\,(|d_n|\cos\theta_n - |y_n|)$$

$$\begin{aligned} |\mu_\mathrm{l}|^2 = \;& \tfrac{(\mu_\mathrm{m} - \mu_\mathrm{p})(|d_n| - |y_n|)}{|d_n - y_n|^2}\,(2\mu_\mathrm{p}\cos\theta_n + \mu_\mathrm{m} - \mu_\mathrm{p})\,|d_n| - \\ & \tfrac{(\mu_\mathrm{m} - \mu_\mathrm{p})(|d_n| - |y_n|)}{|d_n - y_n|^2}\,(\mu_\mathrm{m} + \mu_\mathrm{p})\,|y_n| + \mu_\mathrm{p}^2 \end{aligned}$$

from where

$$2\Re[\mu_\mathrm{l}] - |\mu_\mathrm{l}|^2 = g(\cos\theta_n)\,|d_n|^2\,|d_n - y_n|^{-2} \quad (5)$$

with

$$\begin{aligned} & g(\cos\theta_n) \\ &= 2\cos\theta_n\left[(1 - \mu_\mathrm{p})\,(\mu_\mathrm{m} - \mu_\mathrm{p}) + (\mu_\mathrm{m} + \mu_\mathrm{p} - \mu_\mathrm{m}\mu_\mathrm{p})\,\tfrac{|y_n|}{|d_n|}\right] + \\ & 2\mu_\mathrm{p} - \mu_\mathrm{p}^2 + 2\mu_\mathrm{m}\tfrac{|y_n|^2}{|d_n|^2} + 2\,(\mu_\mathrm{p} - \mu_\mathrm{m})\,\tfrac{|y_n|}{|d_n|} - \left(\mu_\mathrm{m} - \mu_\mathrm{p} - \mu_\mathrm{m}\tfrac{|y_n|}{|d_n|}\right)^2 \end{aligned}$$

[1]The quaternion-valued stepsize is an extension of the complex-valued stepsize in complex gradient learning algorithms [21]–[23].

being a linear function of $\cos\theta_n \in [-1,1]$ which attains the extrema when $\cos\theta_n$ attains the extrema. Therefore, if $g(1) > 0$ and $g(-1) > 0$, then $g(\cos\theta_n) > 0, \forall\theta_n \in [0,\pi]$, so that $2\Re[\mu_l] - |\mu_l|^2 > 0$ always holds.

We can now deduce the sufficient condition for $g(1) > 0$ and $g(-1) > 0$. It is obvious that $g(1) > 0$ if and only if $0 < \mu_{\mathrm{m}} < 2$. Also note that

$$\begin{aligned} g(-1) = \ & \left(2\mu_{\mathrm{m}} - \mu_{\mathrm{m}}^2\right)\tfrac{|y_n|^2}{|d_n|^2} + 2\left(\mu_{\mathrm{m}}^2 - 2\mu_{\mathrm{m}}\mu_{\mathrm{p}} + 2\mu_{\mathrm{p}}\right)\tfrac{|y_n|}{|d_n|} \\ & + \left(2\mu_{\mathrm{p}} - \mu_{\mathrm{m}}\right)\left(2 + \mu_{\mathrm{m}} - 2\mu_{\mathrm{p}}\right). \end{aligned}$$

When $|y_n| = 0$, it can be proved that $g(-1) > 0$ if $\frac{\mu_{\mathrm{m}}}{2} < \mu_{\mathrm{p}} < 1 + \frac{\mu_{\mathrm{m}}}{2}$. Under the condition $\frac{\mu_{\mathrm{m}}}{2} < \mu_{\mathrm{p}} < 1 + \frac{\mu_{\mathrm{m}}}{2}$, $g(-1)$ increases with $\frac{|y_n|}{|d_n|} \in [0, +\infty)$, so $g(-1) > 0, \forall d_n, y_n$. This yields a sufficient convergence condition for QLMMP:

$$0 < \mu_{\mathrm{m}} < 2, \qquad \tfrac{1}{2}\mu_{\mathrm{m}} < \mu_{\mathrm{p}} < 1 + \tfrac{1}{2}\mu_{\mathrm{m}}. \tag{6}$$

Practically, the stepsizes can exceed the sufficient condition given by (6), as shown in Section IV.

B. Special cases of QLMMP

1) For $\mu_{\mathrm{m}} = \mu_{\mathrm{p}}$, the QLMMP is equivalent to the QLMS with a real-valued stepsize μ_{m}. The sufficient convergence condition in (6) reduces to $0 < \mu_{\mathrm{m}} < 2$, and the steady-state excess MSE (EMSE) is obtained from (15) as $(2-\mu_{\mathrm{m}})^{-1}\mu_{\mathrm{m}}\sigma_\eta^2$.

2) For $\mu_{\mathrm{m}} \neq 0$, $\mu_{\mathrm{p}} = 0$, the QLMMP becomes a magnitude-only filtering algorithm which is equivalent to the QLMS with a real-valued stepsize μ_{m} under the constraint that the estimate y_n always aligns with the desired signal d_n, that is, $\theta_n \equiv 0$, $n = 1, 2, \ldots$. The sufficient convergence condition in (6) then reduces to $0 < \mu_{\mathrm{m}} < 2$, while the steady-state EMSE is obtained from (15) as $(2-\mu_{\mathrm{m}})^{-1}\mu_{\mathrm{m}}\sigma_\eta^2$.

3) For $\mu_{\mathrm{m}} = 0$, $\mu_{\mathrm{p}} \neq 0$, the QLMMP reduces to a phase filtering algorithm for which the following holds:

$$y_{n+1} = y_n + \mu_{\mathrm{p}}\left(d_n - |d_n|\,|y_n|^{-1}\,y_n\right) \tag{7}$$

- If $\mu_{\mathrm{p}} = |y_n|\,|d_n|^{-1}$, phase estimation is completed in one iteration for an arbitrary value of phase difference. This is a fundamental advantage over the set of LMP algorithms introduced in [15], [19], for which the phase change in one iteration must be less than $\pi/2$.
- If $\mu_{\mathrm{p}} < |y_n|\,|d_n|^{-1}$, then $|y_{n+1}| < |y_n|$, and the phase difference θ_n monotonically converges to zero.
- If $\mu_{\mathrm{p}} > |y_n|\,|d_n|^{-1}$, then y_{n+1} and y_n lie on two sides of d_n on the plane shown in Fig. 1, and $|y_{n+1}| > |y_n|$, so that the magnitude of the estimate increases until $\mu_{\mathrm{p}} \leq |y_n|\,|d_n|^{-1}$. This adaptive behaviour indicates the stability of the phase filtering algorithm for all positive μ_{p}.

When $\mu_{\mathrm{m}} = 0$, $\mu_{\mathrm{p}} \neq 0$, and for stationary signals, we obtain the steady-state condition: $\angle y_n \approx \angle d_n$, $|d_n| \approx |d_{n+1}|$, $|y_n| \approx |y_{n+1}|$. Therefore, after each weight update, we can implement a magnitude adjustment to the output by $y_{n+1}^{\mathrm{new}} = |d_n|\,|y_n|^{-1}\,y_{n+1}$, and hence perform the complete estimation of both phase and magnitude of d_{n+1}. This algorithm, termed modified QLMMP (MQLMMP), can be used as an alternative, but with superior stability, to the QLMS.

IV. Simulations

The considered class of QLMS algorithms were used to estimate the frequency of a simulated three-phase power system for which the voltages at time n are given by [24]

$$v_{m,n} = V_n \sin\left(2\pi f T n + \varphi + \frac{2\pi}{3}m - \frac{2\pi}{3}\right), \ m = 1, 2, 3 \tag{8}$$

where V_n is the instantaneous magnitude, φ the initial phase, f the frequency, and T the sampling interval. The three voltages can be combined into a pure quaternion given by [25]

$$\begin{aligned} q_n &= \imath v_{1,n} + \jmath v_{2,n} + \kappa v_{3,n} \\ &= \sqrt{1.5}V_n\left[\zeta\sin(2\pi f T n + \varphi) + \zeta'\cos(2\pi f T n + \varphi)\right] \\ &= \sqrt{1.5}V_n e^{-\zeta''(2\pi f T n + \varphi)}\zeta' \end{aligned} \tag{9}$$

where $\zeta = \sqrt{\frac{2}{3}}\left(\imath - \frac{1}{2}\jmath - \frac{1}{2}\kappa\right), \zeta' = \sqrt{0.5}\left(\jmath - \kappa\right), \zeta'' = \sqrt{\frac{1}{3}}\left(\imath + \jmath + \kappa\right)$ are the imaginary units. From (9), then

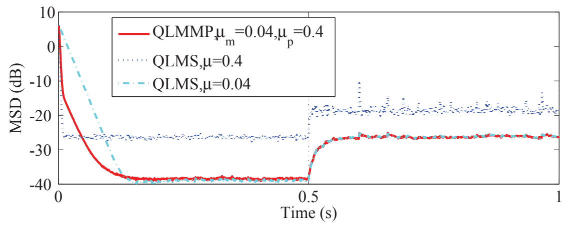
$$q_{n+1} = e^{-\zeta'' 2\pi f T} q_n. \tag{10}$$

The frequency of the three-phase power system can be adaptively calculated through the identification of the system in (10). We used a simulated three-phase power system, for which $f = 50\,\mathrm{Hz}$, $V_n = 1\,\mathrm{V}$, $T = 1\,\mathrm{ms}$, and considered the following two cases of voltage measurements corrupted by different time-varying noise on magnitude and phase.

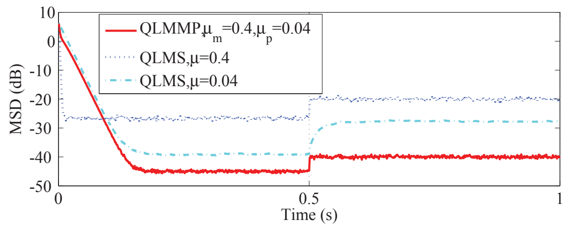
Signal 1: The magnitudes of the three voltages in (8) were corrupted by the same WGN, a_n, with mean 1 and variance changing from 0.01 to 0.04 at 0.5 s. The quaternion-valued measurements of the voltages were additionally subject to additive noise, hence $\sqrt{1.5}a_n V_n e^{-\zeta''(2\pi f T n + \varphi)}\zeta' + \eta_n$, where η_n was quaternion-valued circular WGN with variance 0.0003.

Signal 2: The phases of the three voltages in (8) were corrupted by the same zero-mean WGN, β_n, with variance changing from 0.01 to 0.04 at 0.5 s. The quaternion-valued measurements of the voltages were additionally subject to additive noise, hence $\sqrt{1.5}V_n e^{-\zeta''(2\pi f T n + \varphi + \beta_n)}\zeta' + \eta_n$.

Fig. 2 illustrates the evolution of the mean square deviation (MSD) averaged over 200 simulation runs for three algorithms: 1) normalised QLMMP with $\mu_{\mathrm{m}} = 0.04$ and $\mu_{\mathrm{p}} = 0.4$, 2) normalised QLMS with $\mu = 0.4$, and 3) normalised QLMS with $\mu = 0.04$. As shown in Fig. 2 (a), for Signal 1, the QLMMP achieved a lower steady-state MSD than the QLMS with $\mu = 0.4$. Although the QLMS with $\mu = 0.04$ had a similar steady-state MSD to the QLMMP, its convergence was slower. Fig. 2 (b) indicates that the QLMMP attained the least steady-state MSD among the three algorithms and also converged faster than the QLMS with $\mu = 0.04$. Fig. 3 shows the evolution of the phase error power averaged over 200 simulation runs for two algorithms: 1) normalised QLMMP with $\mu_{\mathrm{m}} = 0$ and $\mu_{\mathrm{p}} = 0.4$, and 2) normalised QLMP with $\mu = 0.4$ [19]. Observe that the QLMMP converged faster than the QLMP in the phase estimation. Fig. 4 shows the evolution of the MSD averaged over 200 simulation runs for the normalised MQLMMP with $\mu_{\mathrm{p}} = 0.4$ introduced in Section III-B3. Figs. 2 and 4 indicate that the accuracy of MQLMMP was lower than that of the QLMMP and QLMS.

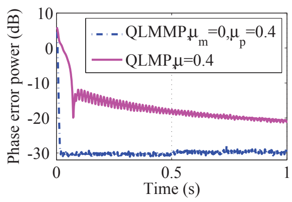


(a) Signal 1

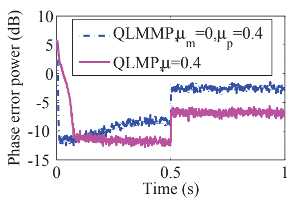


(b) Signal 2

Fig. 2. MSD curves for filtering three-phase voltages using the normalised QLMMP and normalised QLMS.

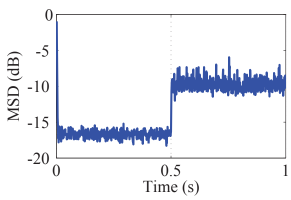


(a) Signal 1

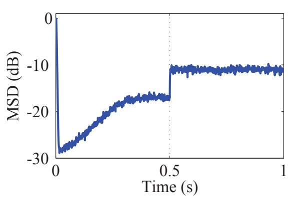


(b) Signal 2

Fig. 3. Phase error power curves for filtering three-phase voltages using the normalised QLMMP and normalised QLMP.



(a) Signal 1



(b) Signal 2

Fig. 4. MSD curves for filtering three-phase voltages using the normalised MQLMMP.

V. Conclusion

We have defined a novel cost function for quaternion adaptive filtering which represents a weighted combination of the magnitude error and phase error. The QLMMP algorithm has been derived based on such a cost function, and theoretical analysis has been conducted to find a range for the stepsizes that guarantees the convergence. The QLMMP has been shown to outperform the QLMS and the QLMP by virtue of the enhanced degrees of freedom in the cost function. The proposed framework is general and can be applied to widely linear and nonlinear processing for quaternion signals [26].

Appendix

The stepsize of the QLMS is conventionally real-valued, but can be alternatively quaternion-valued. Next, we derive the convergence condition and steady-state performance of the normalised QLMS with a quaternion-valued stepsize, $\mu_{\rm l}$. Denote the weight error vector by $\mathbf{v}_n = \mathbf{w}_{\rm o} - \mathbf{w}_n$, then from (1) we obtain

$$\mathbf{v}_{n+1}^* = \mathbf{v}_n^* - \frac{1}{\|\mathbf{x}_n\|_2^2}\mathbf{x}_n\mathbf{x}_n^H\mathbf{v}_n^*\mu_{\rm l}^* - \frac{1}{\|\mathbf{x}_n\|_2^2}\mathbf{x}_n\eta_n^*\mu_{\rm l}^*. \tag{11}$$

The eigendecomposition $E\left\{\mathbf{x}_n\mathbf{x}_n^H\right\} = \mathbf{Q}^H\mathbf{\Lambda}\mathbf{Q}$ enables a coordinate transform given by $\bar{\mathbf{x}}_n = \mathbf{Q}\mathbf{x}_n$, $\bar{\mathbf{v}}_n = \mathbf{Q}\mathbf{v}_n^*$. Applying the statistical expectation operator to (11) yields

$$E\left\{\bar{\mathbf{v}}_{n+1}\right\} = E\left\{\bar{\mathbf{v}}_n\right\}\left(\mathbf{I} - E\left\{\|\bar{\mathbf{x}}_n\|_2^{-2}\,\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^H\right\}\mu_{\rm l}\right). \tag{12}$$

Thus, the sufficient condition for the mean convergence is $\left|1 - \lambda\left(E\left\{\|\bar{\mathbf{x}}_n\|_2^{-2}\,\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^H\right\}\right)\mu_{\rm l}\right| < 1$, which implies

$$2\Re\left[\mu_{\rm l}\right] - \left|\mu_{\rm l}\right|^2\lambda\left(E\left\{\|\bar{\mathbf{x}}_n\|_2^{-2}\,\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^H\right\}\right) > 0 \tag{13}$$

where $\lambda\left(E\left\{\|\bar{\mathbf{x}}_n\|_2^{-2}\,\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^H\right\}\right)$ represents any eigenvalue of the matrix $E\left\{\|\bar{\mathbf{x}}_n\|_2^{-2}\,\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^H\right\}$. It can be proved that $0 < \lambda\left(E\left\{\|\bar{\mathbf{x}}_n\|_2^{-2}\,\bar{\mathbf{x}}_n\bar{\mathbf{x}}_n^H\right\}\right) \leq 1$, and so the expression in (13) can be tightened to $2\Re\left[\mu_{\rm l}\right] - \left|\mu_{\rm l}\right|^2 > 0$.

To analyse the convergence in the mean square sense, from (11) we calculate the variance of the weight error vector as

$$\begin{aligned}\|\mathbf{v}_{n+1}\|_2^2 = \mathbf{v}_{n+1}^T\mathbf{v}_{n+1}^* = &\left(\left|\mu_{\rm l}\right|^2 - 2\Re\left[\mu_{\rm l}\right]\right)\|\mathbf{x}_n\|_2^{-2}\,\mathbf{v}_n^T\mathbf{x}_n\mathbf{x}_n^H\mathbf{v}_n^* + \\ &\mathbf{v}_n^T\mathbf{v}_n^* + \left|\mu_{\rm l}\right|^2\sigma_\eta^2\,\|\mathbf{x}_n\|_2^{-2} + \text{cross terms}.\end{aligned}$$

Upon taking statistical expectations of both sides of the above equation, the mean square deviation (MSD) becomes

$$\begin{aligned}E\left\{\|\mathbf{v}_{n+1}\|_2^2\right\} &= E\left\{\mathbf{v}_n^T\mathbf{F}\mathbf{v}_n^*\right\} + \left|\mu_{\rm l}\right|^2\sigma_\eta^2 E\left\{\|\mathbf{x}_n\|_2^{-2}\right\} \\ \mathbf{F} &= \mathbf{I} + \left(\left|\mu_{\rm l}\right|^2 - 2\Re\left[\mu_{\rm l}\right]\right)E\left\{\|\mathbf{x}_n\|_2^{-2}\,\mathbf{x}_n\mathbf{x}_n^H\right\}.\end{aligned} \tag{14}$$

The recursion in (14) converges if and only if all eigenvalues of $\mathbf{F}$ are within $(-1, 1)$ [27], which is equivalent to

$$-2 < \left(\left|\mu_{\rm l}\right|^2 - 2\Re\left[\mu_{\rm l}\right]\right)\lambda\left(E\left\{\|\mathbf{x}_n\|_2^{-2}\,\mathbf{x}_n\mathbf{x}_n^H\right\}\right) < 0.$$

It can be proved that $0 < \lambda\left(E\left\{\|\mathbf{x}_n\|_2^{-2}\,\mathbf{x}_n\mathbf{x}_n^H\right\}\right) \leq 1$, and so the above inequality can be tightened to $2\Re\left[\mu_{\rm l}\right] - \left|\mu_{\rm l}\right|^2 > 0$.

This establishes that the normalised QLMS converges in the mean and mean square sense for $2\Re\left[\mu_{\rm l}\right] - \left|\mu_{\rm l}\right|^2 > 0$.

Note that (14) yields

$$\begin{aligned}E\left\{\|\mathbf{v}_{n+1}\|_2^2\right\} = \; &E\left\{\|\mathbf{v}_n\|_2^2\right\} + \left|\mu_{\rm l}\right|^2\sigma_\eta^2 E\left\{\|\mathbf{x}_n\|_2^{-2}\right\} \\ &+ \left(\left|\mu_{\rm l}\right|^2 - 2\Re\left[\mu_{\rm l}\right]\right)E\left\{\|\mathbf{x}_n\|_2^{-2}\left|e_n^a\right|^2\right\}\end{aligned}$$

where $e_n^a = \mathbf{v}_n^T\mathbf{x}_n$ is the *a priori* error. At the steady state, $\lim_{n\to\infty} E\left\{\|\mathbf{v}_{n+1}\|_2^2\right\} = \lim_{n\to\infty} E\left\{\|\mathbf{v}_n\|_2^2\right\}$, and we make a further assumption that $\|\mathbf{x}_n\|_2^2$ is statistically independent from $\left|e_n^a\right|^2$ to obtain the steady-state excess MSE (EMSE) in the form

$$\lim_{n\to\infty} E\left\{\left|e_n^a\right|^2\right\} = \left|\mu_{\rm l}\right|^2\sigma_\eta^2\left(2\Re\left[\mu_{\rm l}\right] - \left|\mu_{\rm l}\right|^2\right)^{-1}. \tag{15}$$

REFERENCES

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

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

[3] D. Xu, C. Jahanchahi, C. Cheong Took, and D. P. Mandic, “Enabling quaternion derivatives: The generalized HR calculus,” *Royal Society Open Science*, vol. 2, no. 8, p. 150255, 2015.

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

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

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

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

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

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

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

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

[12] P. J. Schreier and L. L. Scharf, *Statistical signal processing of complex-valued data*. Cambridge, UK: Cambridge Univ. Press, 2010.

[13] C. Cheong Took and D. P. Mandic, “The quaternion LMS algorithm for adaptive filtering of hypercomplex processes,” *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1316–1327, 2009.

[14] J. Navarro-Moreno and J. C. Ruiz-Molina, “Semi-widely linear estimation of $\mathbb{C}^\eta$-proper quaternion random signal vectors under Gaussian and stationary conditions,” *Signal Process.*, vol. 119, pp. 56–66, 2016.

[15] A. Tarighat and A. H. Sayed, “Least mean-phase adaptive filters with application to communications systems,” *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 220–223, 2004.

[16] M. Rupp, “On the separation of channel and frequency offset estimation,” in *Proc. Asilomar Conf. Signals, Systems, Computers*, vol. 2, pp. 1186–1190, 1998.

[17] S. C. Douglas and D. P. Mandic, “The least-mean-magnitude-phase algorithm with applications to communications systems,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 4152–4155, May 2011.

[18] S. C. Douglas and D. P. Mandic, “Stability analysis of the least-mean-magnitude-phase algorithm,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 4935–4939, Mar. 2016.

[19] S. P. Talebi, D. Xu, A. Kuh, and D. P. Mandic, “A quaternion least mean phase adaptive estimator,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 6419–6423, May 2014.

[20] J. Navarro-Moreno, R. Fernandez-Alcala, and J. Ruiz-Molina, “A quaternion widely linear series expansion and its applications,” *IEEE Signal Process. Lett.*, vol. 19, no. 12, pp. 868–871, 2012.

[21] H. Zhang and D. P. Mandic, “Is a complex-valued stepsize advantageous in complex-valued gradient learning algorithms?,” *IEEE Trans. Neural Netw. Learn. Syst.*, vol. PP, no. 99, pp. 1–6, 2015.

[22] D. Nion and L. D. Lathauwer, “An enhanced line search scheme for complex-valued tensor decompositions. Application in DS-CDMA,” *Signal Process.*, vol. 88, no. 3, pp. 749 – 755, 2008.

[23] Y. Chen, D. Han, and L. Qi, “New ALS methods with extrapolating search directions and optimal step size for complex-valued tensor decompositions,” *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5888–5898, 2011.

[24] 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. Proc. Mag.*, vol. 29, no. 5, pp. 44–54, 2012.

[25] S. P. Talebi and D. P. Mandic, “A quaternion frequency estimator for three-phase power systems,” in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, pp. 3956–3960, Apr. 2015.

[26] J. Navarro-Moreno, R. Marı, and J. C. Ruiz-Molina, “A quaternion widely linear model for nonlinear Gaussian estimation,” *IEEE Trans. Signal Process.*, vol. 62, no. 24, pp. 6414–6424, 2014.

[27] T. Y. Al-Naffouri and A. H. Sayed, “Transient analysis of data-normalized adaptive filters,” *IEEE Trans. Signal Process.*, vol. 51, no. 3, pp. 639–652, 2003.