
Article:
Zhang, X., Liu, W., Xu, Y. et al. (2014) Quaternion-valued robust adaptive beamformer for electromagnetic vector-sensor arrays with worst-case constraint. Signal Processing, 104. pp. 274-283. ISSN: 0165-1684

https://doi.org/10.1016/j.sigpro.2014.04.006

Quaternion-valued Robust Adaptive Beamformer for Electromagnetic Vector-sensor Arrays with worst-case Constraint

Xirui Zhang[a,b], Wei Liu[b,*], Yougen Xu[a], Zhiwen Liu[a]

[a] *School of Information and Electronics, Beijing Institute of Technology Beijing, 100081, China*
[b] *Department of Electronic and Electrical Engineering, University of Sheffield Sheffield, S1 3JD, United Kingdom*

Abstract

A robust adaptive beamforming scheme based on two-component electromagnetic (EM) vector-sensor arrays is proposed by extending the well-known worst-case constraint into the quaternion domain. After defining the uncertainty set of the desired signal's quaternionic steering vector, two quaternion-valued constrained minimization problems are derived. We then reformulate them into two real-valued convex quadratic problems, which can be easily solved via the so-called second-order cone (SOC) programming method. It is also demonstrated that the proposed algorithms can be classified as a specific type of the diagonal loading scheme, in which the optimal loading factor is a function of the known level of uncertainty of the desired steering vector. Numerical simulations show that our new method can cope with the steering vector mismatch problem well, and alleviate the finite sample

*Corresponding author: Tel.: +44-114-2225813; fax: +44-114-2225834.
Email addresses: xrzhang@bit.edu.cn (Xirui Zhang), w.liu@sheffield.ac.uk (Wei Liu), yougenxu@bit.edu.cn (Yougen Xu), zwliu@bit.edu.cn (Zhiwen Liu)

size effect to some extent. Besides, the proposed beamformer significantly outperforms the sample matrix inversion minimum variance distortionless response (SMI-MVDR) and the quaternion Capon (Q-Capon) beamformers in all the scenarios studied, and achieves a better performance than the traditional diagonal loading scheme, in the case of smaller sample sizes and higher SNRs.
Keywords: Robust adaptive beamforming, quaternion, electromagnetic vector-sensor array, worst-case constraint, diagonal loading.

---

1. Introduction

Adaptive beamforming with EM vector-sensor arrays can exploit not only the directions of arrival (DOAs) of the impinging signals but also their polarizations. The so-called crossed-dipole and tripole (the earliest EM vector-sensors, also known as the 'polarization diverse antennas') were first introduced into the field of adaptive arrays in [1, 2]. Based on such a system, the adaptive beamforming problem was studied in detail in terms of the output signal-to-interference-plus-noise ratio (SINR) in [3]. Furthermore, it was shown that a 'complete' EM vector-sensor (measuring the six components of an EM field at the same point) with identical electric and magnetic noise power can eliminate the angular grating nulls completely. Moreover, with the analysis in [4], it was concluded that the output-SINR is determined by both DOA and polarization differences of the impinging signals in the context of unequal noise power.

The above methods assume an exactly known steering vector for the desired signal. When the estimation of the steering vector is imprecise, espe-

cially with look direction and sensor position errors, the performance of conventional MVDR beamformer will deteriorate [5]. To enhance its robustness, many methods have been proposed, such as diagonal loading [6, 7, 8, 9], and those based on the optimization of worst-case performance [10, 11, 12, 13]. In particular, the worst-case constrained beamformer (WCCB) can be considered as one specific type of the diagonal loading scheme, where the loading factor is determined based on the known level of uncertainty of the desired signal's steering vector.

Very recently, improved robustness against steering vector mismatch errors has been shown by quaternion formulations [14, 15, 16, 17, 18, 19, 20]. The quaternion-based model of a two-component vector-sensor array was first provided in [14, 15], and a multiple signal classification (MUSIC)-like scheme was applied accordingly. In addition, another subspace-based approach – estimation of signal parameters via rotational invariance techniques (ESPRIT), was also extended to the quaternion domain [16], and this method outperforms the conventional ESPRIT, especially in the circumstances of short data length, low signal-to-noise ratio (SNR) and unknown model errors. For three-component EM vector-sensor arrays, bi-quaternion models were introduced accordingly [21, 22, 23]. In adaptive beamforming, the quaternionic version of the conventional MVDR beamformer has been derived with a two-component EM vector-sensor array in [17, 18], where a better performance is obtained in the presence of steering vector mismatch errors. Afterwards, based on the idea of two-way interference and noise cancellation (INC) [19], a quaternionic adaptive beamforming scheme was presented by adopting a combined structure to achieve an improved performance in the context of one

strong coherent interference [20]. However, the aforementioned quaternion-valued adaptive beamformers have not applied any robust criterion or constraint against the steering vector mismatch problem, and the well-known WCCB has not been investigated in the hypercomplex domain yet. Therefore, a novel quaternion-valued robust adaptive beamformer based on the worst-case constraint is proposed here to tackle the steering vector mismatch problem. First, two adaptive algorithms are derived in detail and solved by the SOC programming method after rearrangement of the parameters. Next, the relationship between the new formulation and the classic diagonal loading scheme is studied using a quaternion-valued Lagrange method. Numerical simulation results indicate that our new methods outperform the SMI-MVDR, and the Q-Capon beamformers [5, 17] for all the examples tested, and is superior to the diagonal loading SMI-MVDR (DL-SMI-MVDR) beamformer [6] for some scenarios. Note a conference version of the basic idea has been published in [24]. Compared with [24], more details of the formulations are provided, supported by extensive simulation results. More importantly, we have derived the exact relationship between the traditional diagonal-loading based scheme and our proposed one in the quaternion domain in Section 3.5 and Appendixes B and C.

The rest of this paper is organized as follows. Sec. 2 introduces the quaternion-related definitions and properties, and then gives the quaternion-based signal model for the two-component EM vector-sensor array. Sec. 3 presents the theoretical derivation of the two proposed algorithms. Numerical simulations are provided in Sec. 4, and conclusions are drawn in Sec. 5.

2. Problem Formulation

2.1. Quaternions

A quaternion $q \in \mathbb{H}^1$ ($\mathbb{R}, \mathbb{C}$, and $\mathbb{H}$ denote the sets of real numbers, complex numbers and quaternions, respectively), is defined as [25]

$$q \triangleq q_0 + q_1\imath + q_2\jmath + q_3\kappa\,, \tag{1}$$

where $q_0 \triangleq \mathrm{Re}\{q\} \in \mathbb{R}^1$ is the real component, while $q_1 \triangleq \mathrm{Im}^{(\imath)}\{q\}$, $q_2 \triangleq \mathrm{Im}^{(\jmath)}\{q\}$, and $q_3 \triangleq \mathrm{Im}^{(\kappa)}\{q\} \in \mathbb{R}^1$ are the three imaginary components, with units $\imath$, $\jmath$, and κ satisfying

$$\begin{aligned} \imath\imath = \jmath\jmath = \kappa\kappa = \imath\jmath\kappa = -1\,, \\ \imath\jmath = -\jmath\imath = \kappa;\ \jmath\kappa = -\kappa\jmath = \imath;\ \kappa\imath = -\imath\kappa = \jmath\,. \end{aligned} \tag{2}$$

The conjugate of a quaternion, denoted by q^*, is defined as

$$q^* \triangleq q_0 - q_1\imath - q_2\jmath - q_3\kappa\,. \tag{3}$$

Specifically, a quaternion q with $q_0 = 0$ is referred to as a pure quaternion. In addition, q can be reformulated into the Cayley-Dickson form as

$$q = \underbrace{q_0 + q_2\jmath}_{\triangleq c_0} + \imath\underbrace{(q_1 + q_3\jmath)}_{\triangleq c_1}\,, \tag{4}$$

where c_0 and $c_1 \in \mathbb{C}$. Similarly, for a vector $\mathbf{v} \in \mathbb{H}^{L_1\times 1}$, and a matrix $\mathbf{M} \in \mathbb{H}^{L_1\times L_2}$, we have

$$\mathbf{v} \triangleq \mathbf{v}_0 + \mathbf{v}_1\imath + \mathbf{v}_2\jmath + \mathbf{v}_3\kappa = \underbrace{\mathbf{v}_0 + \mathbf{v}_2\jmath}_{\triangleq \mathbf{c}_0} + \imath\underbrace{(\mathbf{v}_1 + \mathbf{v}_3\jmath)}_{\triangleq \mathbf{c}_1}\,, \tag{5}$$

$$\mathbf{M} \triangleq \mathbf{M}_0 + \mathbf{M}_1\imath + \mathbf{M}_2\jmath + \mathbf{M}_3\kappa = \underbrace{\mathbf{M}_0 + \mathbf{M}_2\jmath}_{\triangleq \mathbf{C}_0} + \imath\underbrace{(\mathbf{M}_1 + \mathbf{M}_3\jmath)}_{\triangleq \mathbf{C}_1}\,, \tag{6}$$

where $\mathbf{v}_l \in \mathbb{R}^{L_1\times 1}$, $\mathbf{M}_l \in \mathbb{R}^{L_1\times L_2}$, $l = 0, ..., 3$, $\mathbf{c}_0$, $\mathbf{c}_1 \in \mathbb{C}^{L_1\times 1}$, and $\mathbf{C}_0$, $\mathbf{C}_1 \in \mathbb{C}^{L_1\times L_2}$. Furthermore, we need to notice that quaternionic multiplications are noncommutative.

Next, several definitions and properties adopted in this paper are introduced to make the following sections readily comprehensible. More details can be found in [26, 27].

Definition 1. The absolute value $|q|$ is defined as

$$|q| = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}\,. \tag{7}$$

For the quaternionic vector in (5), its Euclidean-norm $\|\mathbf{v}\|$ is

$$\|\mathbf{v}\| = \sqrt{\|\mathbf{v}_0\|_{\mathrm{r}}^2 + \|\mathbf{v}_1\|_{\mathrm{r}}^2 + \|\mathbf{v}_2\|_{\mathrm{r}}^2 + \|\mathbf{v}_3\|_{\mathrm{r}}^2}\,, \tag{8}$$

in which '$\|\cdot\|_{\mathrm{r}}$' denotes the Euclidean-norm for a real-valued vector.

Definition 2. Given two vectors $\mathbf{x}$ and $\mathbf{y} \in \mathbb{H}^{L_1\times 1}$ with their l-th elements respectively denoted by x_l and y_l, their inner product $\mathbf{x}^{\triangleleft}\mathbf{y}$ is defined by

$$\mathbf{x}^{\triangleleft}\mathbf{y} = \sum_{l=1}^{L_1} x_l^* y_l\,, \tag{9}$$

where '$\{\cdot\}^*$' and '$\{\cdot\}^{\triangleleft}$' denote the quaternion conjugate and transposition-conjugate operators, respectively.

Definition 3. The conjugate transpose of the quaternionic matrix given in (6) is

$$\mathbf{M}^{\triangleleft} = \mathbf{M}_0^{\mathrm{T}} - \mathbf{M}_1^{\mathrm{T}}\imath - \mathbf{M}_2^{\mathrm{T}}\jmath - \mathbf{M}_3^{\mathrm{T}}\kappa\,, \tag{10}$$

where '$\{\cdot\}^{\mathrm{T}}$' stands for the transposition operator.

Property 1. Given $\mathbf{M} \in \mathbb{H}^{L_1 \times L_2}$ and $\mathbf{v} \in \mathbb{H}^{L_2 \times 1}$, we have

$$\big(\mathbf{M}\mathbf{v}\big)^{\triangleleft} = \mathbf{v}^{\triangleleft}\mathbf{M}^{\triangleleft}. \tag{11}$$

Property 2. For a quaternion-valued conjugate-symmetric matrix $\mathbf{M} \in \mathbb{H}^{L_1 \times L_1}$ and a vector $\mathbf{v} \in \mathbb{H}^{L_1 \times 1}$, we have [26]

$$\mathbf{v}^{\triangleleft}\mathbf{M}\mathbf{v} = \mathbf{v}_1^{\triangleleft}\mathbf{M}\mathbf{v}_1\,, \quad \text{with} \quad \mathbf{v}_1 = \mathbf{v}\cdot \mathrm{e}^{\epsilon\vartheta}\,, \tag{12}$$

where

$$\begin{aligned} \epsilon &= \frac{q_1\imath + q_2\jmath + q_3\kappa}{\sqrt{q_1^2+q_2^2+q_3^2}}\,, \\ \vartheta &= \arctan\left(\frac{\sqrt{q_1^2+q_2^2+q_3^2}}{q_0}\right). \end{aligned} \tag{13}$$

Property 3. Let $a \in \mathbb{H}^1$, $b \in \mathbb{H}^1$, $\mathbf{a} \in \mathbb{H}^{L_1 \times 1}$, and $\mathbf{b} \in \mathbb{H}^{L_1 \times 1}$. Then

$$\mathrm{Re}\{a\} \geq -|a|\,, \tag{14}$$

$$|a+b| \geq |a| - |b|\,, \tag{15}$$

$$|\mathbf{a}^{\triangleleft}\mathbf{b}| \leq \|\mathbf{a}\|\|\mathbf{b}\|\,, \tag{16}$$

and the equalities hold when a is a non-positive real number for (14), $b = -\varrho a$ with $\varrho \in [0,1]$ for (15), and $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent or both zero vectors for (16). Detailed proofs are provided in Appendix A.

2.2. Quaternion-valued Signal Model

Consider a linear array consisting of N crossed-dipoles (the typical two-component EM vector-sensor) located along the $y-$axis with their position vectors denoted by $\mathbf{r}_n \in \mathbb{R}^{3\times 1}$, $n = 1, 2, ..., N$. As shown in Fig. C.1, the

two components of each crossed-dipole are parallel to $x-$ and $y-$axes, respectively. Suppose there are M uncorrelated narrow-band far-field signals $\{s_m(t)\}_{m=1}^{M}$ impinging upon the array from directions $\{\theta_m, \phi_m\}_{m=1}^{M}$, where $\theta_m \in [0, 2\pi)$ is the azimuth-angle, and $\phi_m \in [0, \pi]$ is the elevation-angle. All the incident signals have the same wavelength λ_0. Then, the spatial steering vector for the mth signal can be expressed as

$$\mathbf{a}_{\mathrm{s},m} = \left[e^{-\jmath\frac{2\pi}{\lambda_0}\mathbf{r}_1^{\mathrm{T}}\boldsymbol{\epsilon}_m}, e^{-\jmath\frac{2\pi}{\lambda_0}\mathbf{r}_2^{\mathrm{T}}\boldsymbol{\epsilon}_m}, \ldots, e^{-\jmath\frac{2\pi}{\lambda_0}\mathbf{r}_N^{\mathrm{T}}\boldsymbol{\epsilon}_m}\right]^{\mathrm{T}}, \tag{17}$$

where

$$\boldsymbol{\epsilon}_m = -[\cos\theta_m \sin\phi_m, \sin\theta_m \sin\phi_m, \cos\phi_m]^{\mathrm{T}}, \tag{18}$$

is the propagation vector of the mth incident signal.

For a crossed-dipole, the spatial-polarization coherent vector of the mth signal with auxiliary polarization angle $\gamma_m \in [0, \pi/2]$ and polarization phase difference $\eta_m \in [-\pi, \pi)$, can be written as

$$\mathbf{a}_{\mathrm{p},m} = \begin{bmatrix} -\sin\theta_m \cos\gamma_m + \cos\theta_m \cos\phi_m \sin\gamma_m e^{\jmath\eta_m} \\ \cos\theta_m \cos\gamma_m + \sin\theta_m \cos\phi_m \sin\gamma_m e^{\jmath\eta_m} \end{bmatrix} = \begin{bmatrix} a_{\mathrm{p},m}^x \\ a_{\mathrm{p},m}^y \end{bmatrix}. \tag{19}$$

Now, we divide the array into two subarrays: one is composed of all the dipoles pointing along the $x-$axis, while the other includes all the dipoles pointing along the $y-$axis. Then, their steering vectors $\mathbf{a}_{x,m}$ and $\mathbf{a}_{y,m} \in \mathbb{C}^{N\times 1}$ for the mth signal are

$$\mathbf{a}_{x,m} = a_{\mathrm{p},m}^x \cdot \mathbf{a}_{\mathrm{s},m}, \quad \mathbf{a}_{y,m} = a_{\mathrm{p},m}^y \cdot \mathbf{a}_{\mathrm{s},m}. \tag{20}$$

The outputs of these two subarrays can then be written as

$$\mathbf{x}(t) = \sum_{m=1}^{M} \mathbf{a}_{x,m} s_m(t) + \mathbf{n}_x(t)\,, \tag{21}$$

$$\mathbf{y}(t) = \sum_{m=1}^{M} \mathbf{a}_{y,m} s_m(t) + \mathbf{n}_y(t)\,, \tag{22}$$

where $\mathbf{n}_x(t)$ and $\mathbf{n}_y(t) \in \mathbb{C}^{N\times 1}$ denote the corresponding additive white Gaussian noise vectors.

Thus, the quaternion-valued output vector $\mathbf{q}(t) \in \mathbb{H}^{N\times 1}$ of the crossed-dipole-based linear array can be defined as

$$\mathbf{q}(t) = \mathbf{x}(t) + \imath\mathbf{y}(t) = \sum_{m=1}^{M} \mathbf{a}_m s_m(t) + \mathbf{n}(t)\,, \tag{23}$$

where $\mathbf{a}_m \triangleq \mathbf{a}_{x,m} + \imath\mathbf{a}_{y,m} \in \mathbb{H}^{N\times 1}$ is the quaternion-valued steering vector, and $\mathbf{n}(t) \triangleq \mathbf{n}_x(t) + \imath\mathbf{n}_y(t) \in \mathbb{H}^{N\times 1}$ is the quaternion-valued noise vector. Given array output data at K distinct snapshots, i.e. $\{\mathbf{q}(t_k)\}_{k=1}^{K}$, the quaternion-valued sample covariance matrix can be evaluated by

$$\hat{\mathbf{R}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{q}(t_k)\mathbf{q}^{\triangleleft}(t_k)\,. \tag{24}$$

3. The Proposed Quaternion-valued Beamformer with Worst-Case Constraint

3.1. Steering Vector Model

Assume that one of the M incident array signals is the desired one and its presumed quaternionic steering vector is denoted as $\bar{\mathbf{a}}_{\mathrm{d}} \in \mathbb{H}^{N\times 1}$. With steering vector mismatch, there will be a non-zero quaternion-valued error vector $\mathbf{e} \in \mathbb{H}^{N\times 1}$ between $\bar{\mathbf{a}}_{\mathrm{d}}$ and the actual steering vector $\mathbf{a}_{\mathrm{d}}$, i.e.,

$$\mathbf{a}_{\mathrm{d}} = \bar{\mathbf{a}}_{\mathrm{d}} + \mathbf{e}. \tag{25}$$

We assume that its norm is bounded by a real positive constant ε, i.e. $\|\mathbf{e}\| \leq \varepsilon$. Then, the actual steering vector $\mathbf{a}_{\mathrm{d}}$ can be modelled as belonging to a steering vector set $\mathcal{A}$ defined by

$$\mathcal{A} \triangleq \left\{ \mathbf{a}_{\mathrm{d}} | \mathbf{a}_{\mathrm{d}} = \bar{\mathbf{a}}_{\mathrm{d}} + \mathbf{e} \,, \|\mathbf{e}\| \leq \varepsilon \right\}. \tag{26}$$

From (26), we can see that $\mathcal{A}$ is a spherical set where $\bar{\mathbf{a}}_{\mathrm{d}}$ is in the center, and $\mathbf{a}_{\mathrm{d}}$ can be any vector in $\mathcal{A}$.

3.2. Quaternion-valued Worst-Case Constrained Algorithm 1

Since $\mathbf{a}_{\mathrm{d}}$ can be any vector in the spherical set $\mathcal{A}$, in order to have a robust response to the desired signal, we can impose the following constraint to the weight vector $\mathbf{w} \in \mathbb{H}^{N \times 1}$

$$\min_{\mathbf{a}_{\mathrm{d}} \in \mathcal{A}} |\mathbf{w}^{\lhd} \mathbf{a}_{\mathrm{d}}| \geq 1 \,, \tag{27}$$

which is referred to as the quaternionic worst-case constraint. Under such a constraint, the magnitude of the array response for all the steering vectors in set $\mathcal{A}$ is constrained to be greater than unity.

By adopting (27), a novel robust adaptive beamformer within the quaternionic framework, named quaternion-valued worst-case constrained beamformer (QWCCB), can be formulated as follows

$$\min_{\mathbf{w}} \mathbf{w}^{\lhd} \hat{\mathbf{R}} \mathbf{w} \quad \text{s.t.} \quad \min_{\mathbf{a}_{\mathrm{d}} \in \mathcal{A}} |\mathbf{w}^{\lhd} \mathbf{a}_{\mathrm{d}}| \geq 1 \,, \tag{28}$$

where $\hat{\mathbf{R}}$ is the sample quaternionic covariance matrix in (24). In the next, we will reformulate the problem in (28), so that it can be solved by SOC programming based method.

Firstly, using the triangle inequality property in (15) along with (25), we have

$$|\mathbf{w}^{\triangleleft}\mathbf{a}_{\mathrm{d}}| = |\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}} + \mathbf{w}^{\triangleleft}\mathbf{e}| \geq |\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}| - |\mathbf{w}^{\triangleleft}\mathbf{e}|\,. \tag{29}$$

Applying the Cauchy-Schwarz inequality in (16) to $|\mathbf{w}^{\triangleleft}\mathbf{e}|$ and with $\|\mathbf{e}\| \leq \varepsilon$, we further have

$$|\mathbf{w}^{\triangleleft}\mathbf{e}| \leq \|\mathbf{w}\|\|\mathbf{e}\| \leq \varepsilon\|\mathbf{w}\|\,. \tag{30}$$

Combining (29) and (30) leads to

$$|\mathbf{w}^{\triangleleft}\mathbf{a}_{\mathrm{d}}| \geq |\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}| - \varepsilon\|\mathbf{w}\|\,. \tag{31}$$

Based on the Theorems 1 and 2 in Appendix A,

$$|\mathbf{w}^{\triangleleft}\mathbf{a}_{\mathrm{d}}| = |\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}| - \varepsilon\|\mathbf{w}\| \tag{32}$$

can be satisfied when 1) ε is small enough to make $|\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}| - \varepsilon\|\mathbf{w}\| > 0$; 2) $\mathbf{e} = -\mathbf{w}\|\mathbf{w}\|^{-1}\varepsilon e^{\epsilon_1\vartheta_1}$, with $\epsilon_1 = \mathrm{Im}\{\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}\}|\mathrm{Im}\{\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}\}|^{-1}$ and $\vartheta_1 = \arctan\big(|\mathrm{Im}\{\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}\}|\cdot \mathrm{Re}^{-1}\{\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}\}\big)$, where $\mathrm{Im}\{\cdot\} \triangleq \mathrm{Im}^{(\imath)}\{\cdot\}\imath + \mathrm{Im}^{(\jmath)}\{\cdot\}\jmath + \mathrm{Im}^{(\kappa)}\{\cdot\}\kappa$. Thus, we can obtain

$$\min_{\mathbf{a}_{\mathrm{d}}\in\mathcal{A}} |\mathbf{w}^{\triangleleft}\mathbf{a}_{\mathrm{d}}| = |\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}| - \varepsilon\|\mathbf{w}\|\,. \tag{33}$$

As a result, the constrained minimization problem in (28) can be transformed into

$$\min_{\mathbf{w}} \mathbf{w}^{\triangleleft}\hat{\mathbf{R}}\mathbf{w} \quad \text{s.t.} \quad |\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}| \geq 1 + \varepsilon\|\mathbf{w}\|\,. \tag{34}$$

However, due to the absolute operation in the constraint, (34) is still a nonconvex problem. According to Property 2, the beamformer's output power $\mathbf{w}^{\triangleleft}\hat{\mathbf{R}}\mathbf{w}$ would not change if the quaternionic vector $\mathbf{w}$ undergoes an arbitrary phase shift. For a given level of $\mathbf{w}^{\triangleleft}\hat{\mathbf{R}}\mathbf{w}$, we can shift the phase of $\mathbf{w}$

without affecting $|\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}|$. Multiplying the weight vector $\mathbf{w}$ by an appropriate phase factor, we can always make $\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}$ a real value, whilst keeping the output power unchanged. Then the constraint in (34) can be rewritten as

$$\begin{aligned}\mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\} \geq 1+\varepsilon\|\mathbf{w}\|\,,\ \mathrm{Im}^{(\imath)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}=0\,,\\ \mathrm{Im}^{(\jmath)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}=0\,,\ \mathrm{Im}^{(\kappa)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}=0\,.\end{aligned} \tag{35}$$

Thus, the constrained minimization problem in (34) can be reformulated into a convex quadratic problem as follows

$$\begin{aligned}\min_{\mathbf{w}}\mathbf{w}^{\lhd}\hat{\mathbf{R}}\mathbf{w}\quad\text{s.t.}\quad &\mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\} \geq 1+\varepsilon\|\mathbf{w}\|\,,\ \mathrm{Im}^{(\imath)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}=0\,,\\ &\mathrm{Im}^{(\jmath)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}=0\,,\ \mathrm{Im}^{(\kappa)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}=0\,.\end{aligned} \tag{36}$$

We refer to the above formulation as the quaternion-valued worst-case constrained beamformer 1 (QWCCB-1).

3.3. Quaternion-valued Worst-Case Constrained Algorithm 2

Following the argument after (34), in the second algorithm, instead of imposing the constraint on the absolute value of $\mathbf{w}^{\lhd}\mathbf{a}_{\mathrm{d}}$, the worst-case constraint is imposed on the real component of $\mathbf{w}^{\lhd}\mathbf{a}_{\mathrm{d}}$, given as

$$\min_{\mathbf{w}}\mathbf{w}^{\lhd}\hat{\mathbf{R}}\mathbf{w}\quad\text{s.t.}\quad\min_{\mathbf{a}_{\mathrm{d}}\in\mathcal{A}}\mathrm{Re}\{\mathbf{w}^{\lhd}\mathbf{a}_{\mathrm{d}}\}\geq 1\,. \tag{37}$$

Using the inequality in (14) along with (25), we have

$$\begin{aligned}\mathrm{Re}\{\mathbf{w}^{\lhd}\mathbf{a}_{\mathrm{d}}\} &= \mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}+\mathrm{Re}\{\mathbf{w}^{\lhd}\mathbf{e}\}\\ &\geq \mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}-|\mathbf{w}^{\lhd}\mathbf{e}|\geq\mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}-\varepsilon\|\mathbf{w}\|\,.\end{aligned} \tag{38}$$

Using Lemma 1 and Theorem 2 in Appendix A, it is easy to verify that

$$\mathrm{Re}\{\mathbf{w}^{\lhd}\mathbf{a}_{\mathrm{d}}\}=\mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}-\varepsilon\|\mathbf{w}\|\,, \tag{39}$$

if 1) ε is small enough to ensure $\mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}-\varepsilon\|\mathbf{w}\|>0$; if 2) $\mathbf{e}=-\varepsilon\mathbf{w}\|\mathbf{w}\|^{-1}$. Then we can conclude that

$$\min_{\mathbf{a}_{\mathrm{d}}\in\mathcal{A}}\mathrm{Re}\{\mathbf{w}^{\lhd}\mathbf{a}_{\mathrm{d}}\}=\mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}-\varepsilon\|\mathbf{w}\|\,. \tag{40}$$

Consequently, the constraint in (37) can be replaced by

$$\mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}-\varepsilon\|\mathbf{w}\|\geq 1\,. \tag{41}$$

The problem (37) can therefore be reformulated into

$$\min_{\mathbf{w}}\mathbf{w}^{\lhd}\hat{\mathbf{R}}\mathbf{w}\quad\text{s.t.}\quad\mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\}\geq 1+\varepsilon\|\mathbf{w}\|\,, \tag{42}$$

which is referred to as the quaternion-valued worst-case constrained beamformer 2 (QWCCB-2).

3.4. SOC Implementation of Q-WCCB-1 and Q-WCCB-2

The constrained minimization problems in (36) and (42) can be solved by the SOC programming method. A SOC program is a convex optimization problem with the following form

$$\min\mathbf{f}^{\mathrm{T}}\mathbf{x}\quad\text{s.t.}\quad\|\mathbf{A}_i\mathbf{x}+\mathbf{b}_i\|\leq\mathbf{c}_i^{\mathrm{T}}\mathbf{x}+d_i, i=1,...,I\,, \tag{43}$$

where $\mathbf{x}\in\mathbb{R}^{N\times 1}$ is the optimization variable, $\mathbf{f}\in\mathbb{R}^{N\times 1}$ denotes the known parameter vector, $\mathbf{A}_i\in\mathbb{R}^{N_i\times N}$, $\mathbf{b}_i\in\mathbb{R}^{N_i\times 1}$, $\mathbf{c}_i\in\mathbb{R}^{N_i\times 1}$, $d_i\in\mathbb{R}$, and I is the number of constraints.

Applying the Cholesky decomposition to $\hat{\mathbf{R}}$ [28], we have

$$\hat{\mathbf{R}}=\mathbf{Q}^{\lhd}\mathbf{Q}\,, \tag{44}$$

where $\mathbf{Q} \in \mathbb{H}^{N\times N}$ is an upper triangular quaternionic matrix. Then the array's output power $\mathbf{w}^{\lhd}\hat{\mathbf{R}}\mathbf{w}$ can be rewritten as

$$\mathbf{w}^{\lhd}\hat{\mathbf{R}}\mathbf{w} = \mathbf{w}^{\lhd}\mathbf{Q}^{\lhd}\mathbf{Q}\mathbf{w} = (\mathbf{Q}\mathbf{w})^{\lhd}(\mathbf{Q}\mathbf{w}) = \|\mathbf{Q}\mathbf{w}\|^2 . \tag{45}$$

Now, by adopting a new nonnegative scalar variable ξ and a new constraint $\|\mathbf{Q}\mathbf{w}\| \le \xi$, the constrained minimization problems in (36) and (42) can be respectively transformed into

$$\begin{aligned} \min_{\mathbf{w},\xi} \xi \quad \text{s.t.} \qquad & \|\mathbf{Q}\mathbf{w}\| \le \xi\,,\ \varepsilon\|\mathbf{w}\| \le \mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\} - 1\,, \\ & \mathrm{Im}^{(\imath)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\} = 0\,,\ \mathrm{Im}^{(\jmath)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\} = 0\,, \\ & \mathrm{Im}^{(\kappa)}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\} = 0\,, \end{aligned} \tag{46}$$

and

$$\min_{\mathbf{w},\xi} \xi \quad \text{s.t.} \qquad \|\mathbf{Q}\mathbf{w}\| \le \xi\,,\ \varepsilon\|\mathbf{w}\| \le \mathrm{Re}\{\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\} - 1\,. \tag{47}$$

Notice that the elements of $\mathbf{Q}$, $\mathbf{w}$, and $\bar{\mathbf{a}}_{\mathrm{d}}$ are quaternions. To facilitate the solution of (46) and (47), we need to convert them into real-valued forms. First of all, $\mathbf{Q}$, $\mathbf{w}$, and $\bar{\mathbf{a}}_{\mathrm{d}}$ need to be written into the following forms

$$\begin{aligned} \mathbf{Q} &\triangleq \mathbf{Q}_0 + \mathbf{Q}_1\imath + \mathbf{Q}_2\jmath + \mathbf{Q}_3\kappa\,, \\ \mathbf{w} &\triangleq \mathbf{w}_0 + \mathbf{w}_1\imath + \mathbf{w}_2\jmath + \mathbf{w}_3\kappa\,, \\ \bar{\mathbf{a}}_{\mathrm{d}} &\triangleq \bar{\mathbf{a}}_{\mathrm{d},0} + \bar{\mathbf{a}}_{\mathrm{d},1}\imath + \bar{\mathbf{a}}_{\mathrm{d},2}\jmath + \bar{\mathbf{a}}_{\mathrm{d},3}\kappa\,, \end{aligned} \tag{48}$$

where $\mathbf{Q}_l \in \mathbb{R}^{N\times N}$, $\mathbf{w}_l \in \mathbb{R}^{N\times 1}$ and $\bar{\mathbf{a}}_{\mathrm{d},l} \in \mathbb{R}^{N\times 1}$, $l = 0, 1, 2, 3$. Then, we

further define one real-valued matrix and five real-valued vectors as follows

$$
\begin{aligned}
\breve{\mathbf{Q}} &\triangleq \begin{bmatrix} \mathbf{Q}_0 & -\mathbf{Q}_1 & -\mathbf{Q}_2 & -\mathbf{Q}_3 \\ \mathbf{Q}_1 & \mathbf{Q}_0 & -\mathbf{Q}_3 & \mathbf{Q}_2 \\ \mathbf{Q}_2 & \mathbf{Q}_3 & \mathbf{Q}_0 & -\mathbf{Q}_1 \\ \mathbf{Q}_3 & -\mathbf{Q}_2 & \mathbf{Q}_1 & \mathbf{Q}_0 \end{bmatrix}, \\
\breve{\mathbf{w}} &\triangleq \left[\mathbf{w}_0^{\mathrm{T}}, \mathbf{w}_1^{\mathrm{T}}, \mathbf{w}_2^{\mathrm{T}}, \mathbf{w}_3^{\mathrm{T}}\right]^{\mathrm{T}}, \\
\bar{\mathbf{a}}_{\mathrm{d}}^{(1)} &\triangleq \left[\bar{\mathbf{a}}_{\mathrm{d},0}^{\mathrm{T}}, \bar{\mathbf{a}}_{\mathrm{d},1}^{\mathrm{T}}, \bar{\mathbf{a}}_{\mathrm{d},2}^{\mathrm{T}}, \bar{\mathbf{a}}_{\mathrm{d},3}^{\mathrm{T}}\right]^{\mathrm{T}}, \\
\bar{\mathbf{a}}_{\mathrm{d}}^{(2)} &\triangleq \left[\bar{\mathbf{a}}_{\mathrm{d},1}^{\mathrm{T}}, -\bar{\mathbf{a}}_{\mathrm{d},0}^{\mathrm{T}}, -\bar{\mathbf{a}}_{\mathrm{d},3}^{\mathrm{T}}, \bar{\mathbf{a}}_{\mathrm{d},2}^{\mathrm{T}}\right]^{\mathrm{T}}, \\
\bar{\mathbf{a}}_{\mathrm{d}}^{(3)} &\triangleq \left[\bar{\mathbf{a}}_{\mathrm{d},2}^{\mathrm{T}}, \bar{\mathbf{a}}_{\mathrm{d},3}^{\mathrm{T}}, -\bar{\mathbf{a}}_{\mathrm{d},0}^{\mathrm{T}}, -\bar{\mathbf{a}}_{\mathrm{d},1}^{\mathrm{T}}\right]^{\mathrm{T}}, \\
\bar{\mathbf{a}}_{\mathrm{d}}^{(4)} &\triangleq \left[\bar{\mathbf{a}}_{\mathrm{d},3}^{\mathrm{T}}, -\bar{\mathbf{a}}_{\mathrm{d},2}^{\mathrm{T}}, \bar{\mathbf{a}}_{\mathrm{d},1}^{\mathrm{T}}, -\bar{\mathbf{a}}_{\mathrm{d},0}^{\mathrm{T}}\right]^{\mathrm{T}}.
\end{aligned} \tag{49}
$$

Based on the above real-valued vectors and matrix, (46) and (47) can be respectively changed into the SOC forms as

$$
\begin{aligned}
\min_{\breve{\mathbf{w}},\xi} \xi \quad \text{s.t.} \qquad & \|\breve{\mathbf{Q}}\breve{\mathbf{w}}\| \leq \xi,\ \varepsilon\|\breve{\mathbf{w}}\| \leq \breve{\mathbf{w}}^{\mathrm{T}}\bar{\mathbf{a}}_{\mathrm{d}}^{(1)} - 1\,, \\
& \breve{\mathbf{w}}^{\mathrm{T}}\bar{\mathbf{a}}_{\mathrm{d}}^{(2)} = 0\,,\ \breve{\mathbf{w}}^{\mathrm{T}}\bar{\mathbf{a}}_{\mathrm{d}}^{(3)} = 0\,, \\
& \breve{\mathbf{w}}^{\mathrm{T}}\bar{\mathbf{a}}_{\mathrm{d}}^{(4)} = 0\,,
\end{aligned} \tag{50}
$$

and

$$
\min_{\breve{\mathbf{w}},\xi} \xi \quad \text{s.t.} \qquad \|\breve{\mathbf{Q}}\breve{\mathbf{w}}\| \leq \xi\,,\ \varepsilon\|\breve{\mathbf{w}}\| \leq \breve{\mathbf{w}}^{\mathrm{T}}\bar{\mathbf{a}}_{\mathrm{d}}^{(1)} - 1\,. \tag{51}
$$

By solving the above constrained minimization problems, the optimum real-valued weight vector $\breve{\mathbf{w}} \in \mathbb{R}^{4N\times 1}$ is obtained. The optimum quaternion-based weight vector $\mathbf{w}_{\mathrm{QWCCB}} \in \mathbb{H}^{N\times 1}$ can then be obtained by re-arranging the elements of $\breve{\mathbf{w}}$ according to (48) and (49).

3.5. Connection with the Diagonal Loading Scheme

To exploit the relationship between the proposed algorithm and the diagonal loading scheme, we first replace the inequality constraint in (28) with

$$\min_{\mathbf{a}_\mathrm{d}\in\mathcal{A}} |\mathbf{w}^{\lhd}\mathbf{a}_\mathrm{d}| = 1\,, \tag{52}$$

which is equivalent to (27), due to the fact that scaling the optimum weight vector does not affect the output SINR [10]. Using (31), the constrained minimization problem (28) can then be simplified to

$$\min_{\mathbf{w}} \mathbf{w}^{\lhd}\hat{\mathbf{R}}\mathbf{w} \quad \text{s.t.} \quad |\mathbf{w}^{\lhd}\bar{\mathbf{a}}_\mathrm{d}| = \varepsilon\|\mathbf{w}\| + 1\,. \tag{53}$$

Following a similar way after (34), for any given $\mathbf{w}$, we can always rotate it in the quaternion domain without affecting the beamformer's output power, so that the imaginary parts of $\mathbf{w}^{\lhd}\bar{\mathbf{a}}_\mathrm{d}$ are fixed to zeros and the real part is positive. Then (53) can be reformulated into

$$\min_{\mathbf{w}} \mathbf{w}^{\lhd}\hat{\mathbf{R}}\mathbf{w} \quad \text{s.t.} \quad |\mathbf{w}^{\lhd}\bar{\mathbf{a}}_\mathrm{d} - 1|^2 = \varepsilon^2\|\mathbf{w}\|^2\,. \tag{54}$$

Further, we can have the quaternion-valued Lagrange objective function as

$$H(\mathbf{w},\zeta) = \mathbf{w}^{\lhd}\hat{\mathbf{R}}\mathbf{w} + \zeta(\mathbf{w}^{\lhd}\bar{\mathbf{a}}_\mathrm{d}\bar{\mathbf{a}}_\mathrm{d}^{\lhd}\mathbf{w} - \bar{\mathbf{a}}_\mathrm{d}^{\lhd}\mathbf{w} - \mathbf{w}^{\lhd}\bar{\mathbf{a}}_\mathrm{d} - \varepsilon^2\mathbf{w}^{\lhd}\mathbf{w} + 1)\,, \tag{55}$$

where ζ is a Lagrange multiplier. Taking the conjugate-gradient of $H(\mathbf{w},\zeta)$ [29] and setting it to zero, yields (see Appendix B for details)

$$\mathbf{w} = \frac{\hat{\mathbf{R}}_\mathrm{DL}^{-1}\bar{\mathbf{a}}_\mathrm{d}}{\bar{\mathbf{a}}_\mathrm{d}^{\lhd}\hat{\mathbf{R}}_\mathrm{DL}^{-1}\bar{\mathbf{a}}_\mathrm{d} + \alpha}\,, \tag{56}$$

where $\alpha \triangleq \zeta^{-1}$, $\beta \triangleq -\varepsilon^2\zeta$, and $\hat{\mathbf{R}}_\mathrm{DL} \triangleq \hat{\mathbf{R}} + \beta\mathbf{I}$.

As a result, the proposed algorithm can be classified as a quaternion domain diagonal loading scheme. To decide the exact value of the Lagrange multiplier ζ, we decompose the sample covariance matrix as $\hat{\mathbf{R}} = \mathbf{E}\mathbf{\Phi}\mathbf{E}^{\lhd}$, with $\mathbf{E}$ and $\mathbf{\Phi} = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_N)$ denoting the eigenvector matrix and the diagonal matrix consisting of all the corresponding eigenvalues of $\hat{\mathbf{R}}$, respectively. Then,

$$\hat{\mathbf{R}}_{\mathrm{DL}}^{-1} = (\hat{\mathbf{R}} + \beta\mathbf{I})^{-1} = \mathbf{E}(\mathbf{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\,. \tag{57}$$

Substituting (57) into (56) leads to

$$\mathbf{w} = \frac{\mathbf{E}(\mathbf{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}}{\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\mathbf{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} + \alpha}\,, \tag{58}$$

which satisfies the constraint in (54), so that we can obtain (see Appendix C for detailed derivations)

$$\zeta^2\varepsilon^2\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\mathbf{\Phi} - \varepsilon^2\zeta\mathbf{I})^{-2}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} = 1\,. \tag{59}$$

Define $\mathbf{z} \triangleq \mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}$, and let z_n denote its nth element. Then, the following result holds

$$f(\zeta) = \sum_{n=1}^{N} \frac{\zeta^2\varepsilon^2|z_n|^2}{(\delta_n - \varepsilon^2\zeta)^2} = 1\,, \tag{60}$$

which is a monotonically decreasing function for $\zeta \in \mathbb{R}^-$. Besides, $f(0) = 0$ and $\lim_{\zeta\to-\infty} f(\zeta) = \frac{1}{\varepsilon^2}\|\mathbf{z}\|^2 = \frac{1}{\varepsilon^2}\|\bar{\mathbf{a}}_{\mathrm{d}}\|^2$. Consequently, in the case of $\varepsilon \in (0, \|\bar{\mathbf{a}}_{\mathrm{d}}\|)$, i.e. $\lim_{\zeta\to-\infty} f(\zeta) > 1$, we can always find a unique negative real number ζ satisfying $f(\zeta) = 1$, and $\beta \triangleq -\varepsilon^2\zeta > 0$.

Instead of solving the nonlinear equation (60) for the Lagrange multiplier ζ, in this work we have adopted the SOC programming based solution introduced in Sec. 3.4 to implement our proposed algorithms.

Note that all the discussions so far have been based on the assumption that the underlying quaternion-valued signals are proper. Otherwise a so-called widely linear model is needed to fully exploit the information carried by the signals for further improved performance [30, 31, 32], which will be a topic of our future research.

4. Simulations

In our simulations, we consider a uniform linear array (ULA) of 10 crossed-dipoles spaced half a wavelength apart. One desired signal along with two uncorrelated interferences is assumed to impinge upon the array from the $y-z$ plane, i.e. $\theta = 90^\circ$, with a fixed signal-to-interference ratio (SIR) of -20 dB, and the elevation-angles for the interferences are fixed to 30° and 60°, respectively. For each scenario, 200 independent Monte-Carlo trials are used to calculate each simulated value. Five approaches are compared using the same type of array in terms of averaged output SINR: the proposed algorithms QWCCB-1 and QWCCB-2, the Q-Capon beamformer [17], the traditional SMI-MVDR beamformer [5], and the DL-SMI-MVDR beamformer with a loading factor of $10\sigma_{\rm n}^2$ [6]. Additionally, the maximally achievable SINR, denoted by 'OPT-SINR', is also displayed in the simulation results as a benchmark. The MATLAB toolboxes SeDuMi [33] and YALMIP [34] are adopted to calculate the weight vectors of our robust beamformers, where the error constraint $\varepsilon = 3$ for the first two examples, and then in the last simulation, ε is varied from 0.6 to 3. The beampattern $\mathrm{B}(\boldsymbol{\vartheta})$ and the output-SINR

are defined as

$$\mathrm{B}(\boldsymbol{\vartheta}) = 20\log_{10}\frac{|\mathbf{w}^{\lhd}\mathbf{a}(\boldsymbol{\vartheta})|}{\max\{|\mathbf{w}^{\lhd}\mathbf{a}(\boldsymbol{\vartheta})|\}}\,, \tag{61}$$

$$\mathrm{SINR} = 10\log_{10}\frac{\mathbf{w}^{\lhd}\mathbf{R}_{\mathrm{d}}\mathbf{w}}{\mathbf{w}^{\lhd}\mathbf{R}_{\mathrm{i+n}}\mathbf{w}}\,, \tag{62}$$

where $\boldsymbol{\vartheta} \triangleq (\phi, \gamma, \eta)$ is the parameter vector, and

$$\mathbf{R}_{\mathrm{d}} = \sigma_{\mathrm{d}}^2\mathbf{a}(\boldsymbol{\vartheta}_{\mathrm{d}})\mathbf{a}^{\lhd}(\boldsymbol{\vartheta}_{\mathrm{d}})\,, \tag{63}$$

$$\mathbf{R}_{\mathrm{i+n}} = \sum_{m=1}^{2}\sigma_{\mathrm{i},m}^2\mathbf{a}(\boldsymbol{\vartheta}_{\mathrm{i},m})\mathbf{a}^{\lhd}(\boldsymbol{\vartheta}_{\mathrm{i},m}) + \sigma_{\mathrm{n}}^2\mathbf{I}\,, \tag{64}$$

with $\mathbf{a}(\boldsymbol{\vartheta}_{\mathrm{d}})$ and $\{\mathbf{a}(\boldsymbol{\vartheta}_{\mathrm{i},m})\}_{m=1}^2$ being the steering vectors of the desired signal and interferences, respectively, and σ_{d}^2 and $\{\sigma_{\mathrm{i},m}^2\}_{m=1}^2$ denoting their corresponding power.

4.1. Example 1: Accurately Known Desired Steering Vector

In the first example, we investigate a scenario where the steering vector of the desired signal is accurately known as prior information. Although no steering vector mismatch errors are taken into account, the finite number of samples may lead to equivalent array model errors.

The desired signal is assumed to arrive from $\phi_{\mathrm{d}} = 5^\circ$ with polarizations $(\gamma_{\mathrm{d}}, \eta_{\mathrm{d}}) = (15^\circ, 50^\circ)$, and interferences' polarizations are $(\gamma_{\mathrm{i},1}, \eta_{\mathrm{i},1}) = (35^\circ, 80^\circ)$ and $(\gamma_{\mathrm{i},2}, \eta_{\mathrm{i},2}) = (85^\circ, 110^\circ)$. The output-SINRs versus the sample size with a fixed SNR of 0 dB is shown in Fig. C.2, and the output-SINRs versus the input SNR for a fixed sample size $K = 30$ is shown in Fig. C.3. From both results, we can observe that the proposed methods consistently outperform the Q-Capon and SMI-MVDR beamformers in terms of robustness to finite sample effect and convergence rate, and is superior to the well-known

DL-SMI-MVDR beamformer especially in the case of smaller sample sizes (less than 40) and higher SNRs (larger than 2 dB).

4.2. Example 2: Steering Vector Mismatch

In this example, we study the beam pattern and output-SINR performance of the proposed algorithms in the presence of steering vector mismatch errors. First, we give the beam patterns of the proposed beamformers to verify their effectiveness to tackle the steering vector mismatch problem, and then compare them with that of Q-Capon beamformer. To draw the 2-D beam pattern, we assume that the polarizations of the incident signals are identical, and the desired signal is from $\phi_{\mathrm{d}} = 5^{\circ}$. Besides, the mismatch error vector $\|\mathbf{e}\|$ is assumed to have a uniform distribution in the interval of $(0, \varepsilon]$. The input SNR is 0 dB, and the sample size is 100. From the resultant beam patterns shown in Fig. C.4, we can see that the proposed algorithms have effectively formed two nulls towards 30° and 60°, and avoided cancellation of the desired signal. By contrast, the Q-Capon beamformer has caused serious suppression to the desired signal from 5°. In addition, QWCCB-1 and QWCCB-2 have almost the same beampattern.

Next, we show the output-SINR performance of the proposed beamformer. The desired signal's DOA and the polarizations of incident signals are kept the same as in the first example. Fig. C.5 shows the output-SINR versus the sample size with a fixed SNR of 0 dB, and Fig. C.6 gives the output-SINR versus the input SNR with a sample size of 30. Clearly, the two proposed beamformers have achieved a better performance than the SMI-MVDR and Q-Capon beamformers for all the sample sizes and SNRs simulated. The DL-SMI-MVDR beamformer performs better than our algorithms in Fig. C.5 at

all sample sizes. However, when we increase the input SNR to about 6 dB, as shown in Fig. C.6, our method becomes better than the DL-SMI-MVDR beamformer.

4.3. Example 3: Output-SINR Versus ε

In the last example, we consider the output-SINR performance of the QWCCB-1 and QWCCB-2 beamformers versus the error constraint ε, for SNR = 0 dB and $K = 50$. The actual DOAs and polarizations of incidents signals are the same as used in Example 1. The results are shown in Fig. C.7, and we can see that QWCCB-1 is better than QWCCB-2 for small values of ε. When the value of ε is greater than about 1.4, these two quaternion-based beamformers have achieved nearly the same performance. Moreover, when ε is very large (close to 3), their output-SINR begins to drop. One explanation for this is, when ε is too large, the significantly increased uncertainty level leaves much less freedom for the beamformer to effectively suppress the interfering signals. Also considering results from previous simulations, overall, it seems that the first algorithm is preferred to the second one.

5. Conclusion

Based on the well-known worst-case performance constraint, a new class of adaptive beamforming algorithms for two-component EM vector-sensor arrays has been proposed within the hypercomplex framework. The quaternion-valued formulation is transformed into a real-valued convex optimization problem, and solved using the SOC programming method. The relationship between our proposed beamformer and the classic diagonal loading scheme has also been investigated, and it is shown that our derived algorithms can

be considered as a specific diagonal loading technique in the quaternion domain, whose loading factor is dependent on the known level of uncertainty of the desired signal's steering vector. Simulation results have verified the robustness of the proposed beamformer towards both steering vector mismatch errors and finite sample size effect, and shown that it has a better performance compared with another three beamformers in most scenarios tested. Additionally, the two proposed algorithms have achieved a very close performance in our numerical simulations. However, the first one is preferred given its overall superior performance.

Appendix A. Proofs of Property 3

Lemma 1. For a quaternion a, we have

$$\mathrm{Re}\{a\} \geq -|a|\,, \tag{A.1}$$

where the equality holds for the case that a is a non-positive real number.

Proof: Consider the following two cases:

1) If $\mathrm{Re}\{a\} > 0$, we can easily have

$$\mathrm{Re}\{a\} > 0 \geq -|a|\,. \tag{A.2}$$

2) If $\mathrm{Re}\{a\} \leq 0$, then

$$\begin{aligned} \mathrm{Re}\{a\} &= -|\mathrm{Re}\{a\}| \\ &\geq -\sqrt{[\mathrm{Re}\{a\}]^2 + [\mathrm{Im}^{\imath}\{a\}]^2 + [\mathrm{Im}^{\jmath}\{a\}]^2 + [\mathrm{Im}^{\kappa}\{a\}]^2} \\ &= -|a|\,. \end{aligned} \tag{A.3}$$

Combining (A.2) and (A.3), we then have (A.1) and the condition for the equality to hold. □

Lemma 2. Given two quaternions a and b, the following holds

$$|ab^*| = |a||b|\,. \tag{A.4}$$

Proof: Based on the Theorem 2.1 in [27], we have

$$|ab^*| = |b^*a| = |a||b^*| = |a||b|\,, \tag{A.5}$$

which completes the proof. □

Theorem 1. Let a and $b \in \mathbb{H}^1$. Then we can obtain that

$$|a+b| \geq |a| - |b|\,, \tag{A.6}$$

where the equality holds when $b = -\varrho a$ with $\varrho \in [0,1]$.

Proof: Here, we will investigate the following two scenarios.

1) Suppose $|a| < |b|$. Then

$$|a+b| \geq 0 > |a| - |b|\,. \tag{A.7}$$

2) Suppose $|a| \geq |b|$. Based on Lemmas 1 and 2, we have

$$2\mathrm{Re}\{ab^*\} \geq -2|ab^*| = -2|a||b|\,. \tag{A.8}$$

Replacing $2\mathrm{Re}\{ab^*\}$ with $ab^* + ba^*$ and using $aa^* + bb^* = |a|^2 + |b|^2$, yields

$$aa^* + bb^* + ab^* + ba^* \geq |a|^2 + |b|^2 - 2|a||b| = (|a| - |b|)^2\,. \tag{A.9}$$

Moreover,

$$aa^* + bb^* + ab^* + ba^* = (a+b)(a+b)^* = |a+b|^2\,, \tag{A.10}$$

which can be substituted into (A.9) to lead to

$$|a+b|^2 \geq (|a| - |b|)^2\,. \tag{A.11}$$

Since $|a| \geq |b|$, we can finally obtain

$$|a+b| \geq |a| - |b|\,. \tag{A.12}$$

With (A.7) and (A.12), (A.6) and the condition satisfying its equality can then be proved. □

Theorem 2. For $\mathbf{a}$ and $\mathbf{b} \in \mathbb{H}^{L_1 \times 1}$, we have

$$|\mathbf{a}^{\triangleleft}\mathbf{b}| \leq \|\mathbf{a}\|\|\mathbf{b}\|\,, \tag{A.13}$$

where the equality holds if $\mathbf{a}$ and $\mathbf{b}$ are linearly dependent or if they are both zero vectors.

Proof: Let $\rho \in \mathbb{H}^1$. Then

$$\big(\mathbf{a}-\mathbf{b}\rho\big)^{\triangleleft}\big(\mathbf{a}-\mathbf{b}\rho\big) = \|\mathbf{a}\|^2 + |\rho|^2\|\mathbf{b}\|^2 - \mathbf{a}^{\triangleleft}\mathbf{b}\rho - \rho^*\mathbf{b}^{\triangleleft}\mathbf{a} \geq 0\,. \tag{A.14}$$

We further substitute $\rho = \|\mathbf{b}\|^{-2}\mathbf{b}^{\triangleleft}\mathbf{a}$ into the above formulation and use $\mathbf{b}^{\triangleleft}\mathbf{a} = (\mathbf{a}^{\triangleleft}\mathbf{b})^{\triangleleft} = (\mathbf{a}^{\triangleleft}\mathbf{b})^*$ (see Theorem 4.1 in [27]) to obtain

$$|\mathbf{a}^{\triangleleft}\mathbf{b}|^2\|\mathbf{b}\|^{-2} \leq \|\mathbf{a}\|^2\,, \tag{A.15}$$

which then yields (A.13). In addition, $\|\mathbf{a}-\mathbf{b}\rho\|^2 = 0$ holds when $\mathbf{a}$ and $\mathbf{b}$ are both zero vectors, or when $\mathbf{a} = \mathbf{b}\rho$ (i.e., they are linearly dependent). □

Appendix B. The Solution for the Quaternion-valued Lagrange Problem

Given a real/quaternion-valued function of the quaternionic variable $w = w_a + w_b\imath + w_c\jmath + w_d\kappa$, denoted as $f(w)$, the direction of its maximum changing rate is given by its conjugate gradient, which is defined as [29]

$$\frac{\partial f(w)}{\partial w^*} \triangleq \frac{1}{4}\left[\frac{\partial f(w)}{\partial w_a} + \imath\frac{\partial f(w)}{\partial w_b} + \jmath\frac{\partial f(w)}{\partial w_c} + \kappa\frac{\partial f(w)}{\partial w_d}\right], \tag{B.1}$$

which can yield $\partial w/\partial w^* = -(1/2)$ and $\partial w^*/\partial w^* = 1$.

In the next, we further consider a real/quaternion-valued function $f(\mathbf{w})$ with the quaternion-valued vector variable $\mathbf{w} \triangleq [w_1, w_2, ..., w_N]^{\mathrm{T}}$. Its partial differential with respect to $\mathbf{w}^*$ can be defined by

$$\frac{\partial f(\mathbf{w})}{\partial \mathbf{w}^*} \triangleq \left[\frac{\partial f(\mathbf{w})}{\partial w_1^*}, \frac{\partial f(\mathbf{w})}{\partial w_2^*}, ..., \frac{\partial f(\mathbf{w})}{\partial w_N^*}\right]^{\mathrm{T}}. \tag{B.2}$$

For a vector function $\mathbf{f}(\mathbf{w}) \triangleq [f_1(\mathbf{w}), f_2(\mathbf{w}), ..., f_N(\mathbf{w})]^{\mathrm{T}}$ with the vector variable $\mathbf{w}$, we further have

$$\frac{\partial \mathbf{f}^{\mathrm{T}}(\mathbf{w})}{\partial \mathbf{w}^*} \triangleq \begin{bmatrix} \frac{\partial f_1(\mathbf{w})}{\partial w_1^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_1^*} & \cdots & \frac{\partial f_N(\mathbf{w})}{\partial w_1^*} \\ \frac{\partial f_1(\mathbf{w})}{\partial w_2^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_2^*} & \cdots & \frac{\partial f_N(\mathbf{w})}{\partial w_2^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1(\mathbf{w})}{\partial w_N^*} & \frac{\partial f_2(\mathbf{w})}{\partial w_N^*} & \cdots & \frac{\partial f_N(\mathbf{w})}{\partial w_N^*} \end{bmatrix}. \tag{B.3}$$

According to (B.1)–(B.3), the following results hold

$$\frac{\partial \mathbf{w}^{\mathrm{T}}}{\partial \mathbf{w}^*} = \begin{bmatrix} \frac{\partial w_1}{\partial w_1^*} & \frac{\partial w_2}{\partial w_1^*} & \cdots & \frac{\partial w_N}{\partial w_1^*} \\ \frac{\partial w_1}{\partial w_2^*} & \frac{\partial w_2}{\partial w_2^*} & \cdots & \frac{\partial w_N}{\partial w_2^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial w_1}{\partial w_N^*} & \frac{\partial w_2}{\partial w_N^*} & \cdots & \frac{\partial w_N}{\partial w_N^*} \end{bmatrix} = -\frac{1}{2}\mathbf{I}_N\,, \tag{B.4}$$

$$\frac{\partial \mathbf{w}^{\lhd}}{\partial \mathbf{w}^*} = \begin{bmatrix} \frac{\partial w_1^*}{\partial w_1^*} & \frac{\partial w_2^*}{\partial w_1^*} & \cdots & \frac{\partial w_N^*}{\partial w_1^*} \\ \frac{\partial w_1^*}{\partial w_2^*} & \frac{\partial w_2^*}{\partial w_2^*} & \cdots & \frac{\partial w_N^*}{\partial w_2^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial w_1^*}{\partial w_N^*} & \frac{\partial w_2^*}{\partial w_N^*} & \cdots & \frac{\partial w_N^*}{\partial w_N^*} \end{bmatrix} = \mathbf{I}_N\,. \tag{B.5}$$

Then we have

$$\begin{aligned}\frac{\partial \mathbf{w}^{\triangleleft}\hat{\mathbf{R}}\mathbf{w}}{\partial \mathbf{w}^{*}} &= \frac{\partial \mathbf{w}^{\triangleleft}}{\partial \mathbf{w}^{*}}\hat{\mathbf{R}}\mathbf{w} + \big(\mathbf{w}^{\triangleleft}\hat{\mathbf{R}}\big)^{\mathrm{T}}\frac{\partial \mathbf{w}^{\mathrm{T}}}{\partial \mathbf{w}^{*}} \\ &= \hat{\mathbf{R}}\mathbf{w} - \frac{1}{2}\big(\mathbf{w}^{\triangleleft}\hat{\mathbf{R}}\big)^{\mathrm{T}}, \end{aligned} \tag{B.6}$$

$$\begin{aligned}\frac{\partial \mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\mathbf{w}}{\partial \mathbf{w}^{*}} &= \frac{\partial \mathbf{w}^{\triangleleft}}{\partial \mathbf{w}^{*}}\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\mathbf{w} + \big(\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\big)^{\mathrm{T}}\frac{\partial \mathbf{w}^{\mathrm{T}}}{\partial \mathbf{w}^{*}} \\ &= \bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\mathbf{w} - \frac{1}{2}\big(\mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\big)^{\mathrm{T}}, \end{aligned} \tag{B.7}$$

$$\frac{\partial \bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\mathbf{w}}{\partial \mathbf{w}^{*}} = \big(\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\big)^{\mathrm{T}}\frac{\partial \mathbf{w}^{\mathrm{T}}}{\partial \mathbf{w}^{*}} = -\frac{1}{2}\bar{\mathbf{a}}_{\mathrm{d}}^{*}, \tag{B.8}$$

$$\frac{\partial \mathbf{w}^{\triangleleft}\bar{\mathbf{a}}_{\mathrm{d}}}{\partial \mathbf{w}^{*}} = \frac{\partial \mathbf{w}^{\triangleleft}}{\partial \mathbf{w}^{*}}\bar{\mathbf{a}}_{\mathrm{d}} = \bar{\mathbf{a}}_{\mathrm{d}}, \tag{B.9}$$

$$\begin{aligned}\frac{\partial \mathbf{w}^{\triangleleft}\mathbf{w}}{\partial \mathbf{w}^{*}} &= \frac{\partial \mathbf{w}^{\triangleleft}}{\partial \mathbf{w}^{*}}\mathbf{w} + (\mathbf{w}^{\triangleleft})^{\mathrm{T}}\frac{\partial \mathbf{w}^{\mathrm{T}}}{\partial \mathbf{w}^{*}} \\ &= \mathbf{w} - \frac{1}{2}\mathbf{w}^{*}. \end{aligned} \tag{B.10}$$

Combining (B.6)–(B.10) along with (55), and using property 1, yields

$$\begin{aligned}\frac{\partial H(\mathbf{w},\zeta)}{\partial \mathbf{w}^{*}} &= \hat{\mathbf{R}}\mathbf{w} - \frac{1}{2}\big(\hat{\mathbf{R}}\mathbf{w}\big)^{*} + \zeta\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\mathbf{w} \\ &\quad - \frac{1}{2}\zeta\big(\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft}\mathbf{w}\big)^{*} + \frac{1}{2}\zeta\bar{\mathbf{a}}_{\mathrm{d}}^{*} - \zeta\bar{\mathbf{a}}_{\mathrm{d}} \\ &\quad - \varepsilon^{2}\zeta\mathbf{w} + \frac{1}{2}\varepsilon^{2}\zeta\mathbf{w}^{*}. \end{aligned} \tag{B.11}$$

By equating (B.11) to zero, we can further have

$$\begin{aligned}&\frac{1}{3}\mathrm{Re}\big\{\big(\hat{\mathbf{R}} + \zeta\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft} - \varepsilon^{2}\zeta\mathbf{I}\big)\mathbf{w}\big\} + \\ &\quad \mathrm{Im}\big\{\big(\hat{\mathbf{R}} + \zeta\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft} - \varepsilon^{2}\zeta\mathbf{I}\big)\mathbf{w}\big\} = \frac{1}{3}\mathrm{Re}\{\zeta\bar{\mathbf{a}}_{\mathrm{d}}\} + \mathrm{Im}\{\zeta\bar{\mathbf{a}}_{\mathrm{d}}\}, \end{aligned} \tag{B.12}$$

where $\mathrm{Im}\{\cdot\} \triangleq \mathrm{Im}^{(\imath)}\{\cdot\}\imath + \mathrm{Im}^{(\jmath)}\{\cdot\}\jmath + \mathrm{Im}^{(\kappa)}\{\cdot\}\kappa$.

From (B.12), we can draw a conclusion as

$$\mathbf{w} = \zeta\Big(\hat{\mathbf{R}} + \zeta\bar{\mathbf{a}}_{\mathrm{d}}\bar{\mathbf{a}}_{\mathrm{d}}^{\triangleleft} - \varepsilon^{2}\zeta\mathbf{I}\Big)^{-1}\bar{\mathbf{a}}_{\mathrm{d}}. \tag{B.13}$$

Using the matrix inversion lemma which can be shown to be applicable in the quaternion domain as long as $(\mathbf{M}_1\mathbf{M}_2)^{-1} = \mathbf{M}_2^{-1}\mathbf{M}_1^{-1}$ holds for invertible quaternion-valued matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ [27], (B.13) can then be reformulated into

$$\mathbf{w} = \frac{\hat{\mathbf{R}}_{\mathrm{DL}}^{-1}\bar{\mathbf{a}}_{\mathrm{d}}}{\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\hat{\mathbf{R}}_{\mathrm{DL}}^{-1}\bar{\mathbf{a}}_{\mathrm{d}} + \alpha}\,, \tag{B.14}$$

where $\alpha \triangleq \zeta^{-1}$, $\beta \triangleq -\varepsilon^2\zeta$, and $\hat{\mathbf{R}}_{\mathrm{DL}} \triangleq \hat{\mathbf{R}} + \beta\mathbf{I}$.

Appendix C. Derivations of Equation (59)

Substituting (58) into the constraint in (54), yields

$$\begin{aligned} |\mathbf{w}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} - 1|^2 &= \left| \frac{\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}}{\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} + \alpha} - 1 \right|^2 \\ &= \frac{\alpha^2}{\left|\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} + \alpha\right|^2}\,, \end{aligned} \tag{C.1}$$

and

$$\varepsilon^2\|\mathbf{w}\|^2 = \varepsilon^2 \left\| \frac{\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}}{\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} + \alpha} \right\|^2 . \tag{C.2}$$

Since $\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} + \alpha$ is a scalar, (C.2) can be rewritten as

$$\varepsilon^2\|\mathbf{w}\|^2 = \frac{\varepsilon^2\|\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\|^2}{\left|\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} + \alpha\right|^2}\,. \tag{C.3}$$

Combining (C.1) and (C.3), leads to

$$\alpha^{-2}\varepsilon^2\|\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\|^2 = 1\,. \tag{C.4}$$

With $\mathbf{E}^{\lhd}\mathbf{E} = \mathbf{I}$, we then have

$$\|\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-1}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\|^2 = \bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\boldsymbol{\Phi} + \beta\mathbf{I})^{-2}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}}\,. \tag{C.5}$$

Finally, by replacing α and β with $\alpha = \zeta^{-1}$ and $\beta = -\varepsilon^2\zeta$, (C.4) can be written as

$$\zeta^2\varepsilon^2\bar{\mathbf{a}}_{\mathrm{d}}^{\lhd}\mathbf{E}(\mathbf{\Phi} - \varepsilon^2\zeta\mathbf{I})^{-2}\mathbf{E}^{\lhd}\bar{\mathbf{a}}_{\mathrm{d}} = 1\,, \tag{C.6}$$

which has been given in (59).

References

[1] R. T. Compton, On the performance of a polarization sensitive adaptive array, IEEE Transactions on Antennas and Propagation 29 (1981) 718–725.

[2] R. T. Compton, The tripole antenna: An adaptive array with full polarization flexibility, IEEE Transactions on Antennas and Propagation 29 (1981) 944–952.

[3] A. Nehorai, K. C. Ho, B. T. G. Tan, Minimum-noise-variance beamformer with an electromagnetic vector sensor, IEEE Transactions on Signal Processing 47 (1999) 601–618.

[4] Y. G. Xu, T. Liu, Z. W. Liu, Output SINR of MV beamformer with one EM vector sensor of and magnetic noise power, in: IEEE International Conference on Signal Processing (ICSP), volume 1, 2004, pp. 419–422.

[5] I. S. Reed, J. D. Mallett, L. E. Brennan, Rapid convergence rate in adaptive arrays, IEEE Transactions on Aerospace and Electronic Systems AES-10 (1974) 853–863.

[6] H. Cox, R. M. Zeskind, M. M. Owen, Robust adaptive beamforming, IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-35 (1987) 1365–1376.

[7] J. Li, P. Stoica (Eds.), Robust Adaptive Beamforming, John Wiley & Sons, Hoboken, New Jersey, 2005.

[8] B. D. Carlson, Covariance matrix estimation errors and diagonal loading in adaptive arrays, IEEE Transactions on Aerospace and Electronic Systems 24 (1988) 397–401.

[9] W. Liu, S. Weiss, Wideband Beamforming: Concepts and Techniques, John Wiley & Sons, Chichester, UK, 2010.

[10] S. A. Vorobyov, A. B. Gershman, Z. Q. Luo, Robust adaptive beamforming using worst-case performance optimization via second-order cone programming, in: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), volume 3, 2002, pp. III–2901–III–2904.

[11] S. A. Vorobyov, A. B. Gershman, Z. Q. Luo, Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem, IEEE Transactions on Signal Processing 51 (2003) 313–324.

[12] L. Yu, W. Liu, R. J. Langley, Novel robust beamformers for coherent interference suppression with DOA estimation errors, IET Microwaves, Antennas and Propagation 4 (2010) 1310–1319.

[13] Y. Zhao, W. Liu, Robust wideband beamforming with frequency response variation constraint subject to arbitrary norm-bounded error, IEEE Transactions on Antennas and Propagation 60 (2012) 2566–2571.

[14] S. Miron, N. Le Bihan, J. I. Mars, High resolution vector-sensor array processing using quaternions, in: IEEE/SP 13th Workshop on Statistical Signal Processing, 2005, pp. 918–923.

[15] S. Miron, N. Le Bihan, J. I. Mars, Quaternion-MUSIC for vector-sensor array processing, IEEE Transactions on Signal Processing 54 (2006) 1218–1229.

[16] X. F. Gong, Y. G. Xu, Z. W. Liu, Quaternion ESPRIT for direction finding with a polarization sentive array, in: IEEE International Conference on Signal Processing (ICSP), 2008, pp. 378–381.

[17] X. M. Gou, Y. G. Xu, Z. W. Liu, X. F. Gong, Quaternion-Capon beamformer using crossed-dipole arrays, in: IEEE International Symposium on Microwave, Antenna, Propagation, and EMC Technologies for Wireless Communications (MAPE), 2011, pp. 34–37.

[18] J.-W. Tao, W.-X. Chang, The MVDR beamformer based on hypercomplex processes, in: IEEE International Conference on Computer Science and Electronics Engineering (ICCSEE), volume 2, 2012, pp. 273–277.

[19] J.-W. Tao, Performance analysis for interference and noise canceller based on hypercomplex and spatio-temporal-polarisation processes, IET Radar, Sonar Navigation 7 (2013) 277–286.

[20] J.-W. Tao, W.-X. Chang, A novel combined beamformer based on hypercomplex processes, IEEE Transactions on Aerospace and Electronic Systems 49 (2013) 1276–1289.

[21] S. Miron, N. Le Bihan, J. I. Mars, High resolution vector-sensor array processing based on biquaternions, in: Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, volume 4, Toulouse, France, 2006, pp. IV–1077–IV–1080.

[22] N. Le Bihan, S. Miron, J. I. Mars, Music algorithm for vector-sensors array using biquaternions, IEEE Transactions on Signal Processing 55 (2007) 4523–4533.

[23] X. F. Gong, Z. W. Liu, Y. G. Xu, Direction finding via biquaternion matrix diagonalization with vector-sensors, Signal Processing 91 (2011) 821–831.

[24] X. R. Zhang, W. Liu, Y. G. Xu, Z. W. Liu, Quaternion-based worst case constrained beamformer based on electromagnetic vector-sensor arrays, in: Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, Vancouver, Canada, 2013, pp. 4149–4153.

[25] W. R. Hamilton, On quaternions, in: Proceeding of the Royal Irish Academy, 1844, pp. 1–16.

[26] N. Le Bihan, J. I. Mars, Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing, Signal Processing 84 (2004) 1177–1199.

[27] F. Zhang, Quaternions and matrices of quaternions, Linear Algebra and its Applications 251 (1997) 21–57.

[28] M. Wang, W. Ma, A structure-preserving algorithm for the quaternion cholesky decomposition, Applied Mathematics and Computation 223 (2013) 354–361.

[29] D. P. Mandic, C. Jahanchahi, C. C. Took, A quaternion gradient operator and its applications 18 (2011) 47–50.

[30] D. P. Mandic, S. L. Goh, Complex Valued Nonlinear Adaptive Filters: Noncircularity, Widely Linear and Neural Models, John Wiley & Sons, UK, 2009.

[31] C. Took, D. Mandic, A quaternion widely linear adaptive filter, IEEE Transactions on Signal Processing 58 (2010) 4427–4431.

[32] J. Via, Properness and widely linear processing of quaternion random vectors, IEEE Trans. Information Theory 56 (2010) 3502–3515.

[33] J. F. Sturm, Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, 1998.

[34] J. Lofberg, YALMIP: A toolbox for modeling and optimization in MATLAB, in: IEEE International Symposium on Computer Aided Control Systems Design, 2004, pp. 284–289.

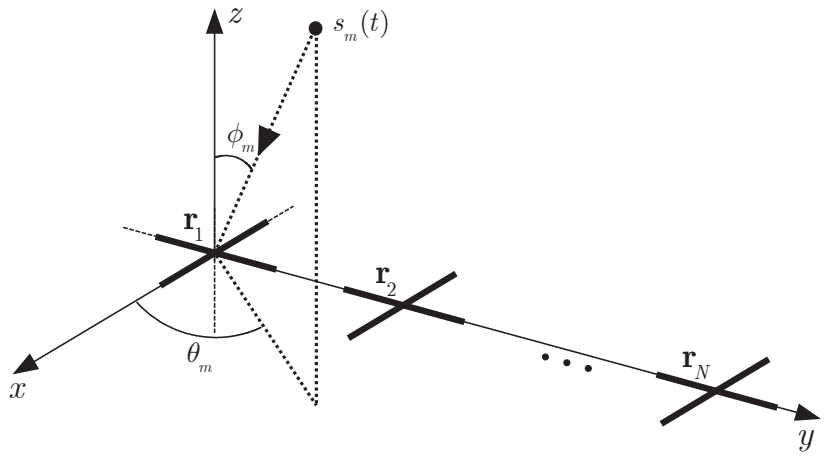

Figure C.1: A linear array of crossed-dipoles: $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N$ denote the position vectors of the crossed-dipoles, θ_m is the azimuth-angle, and ϕ_m is the elevation-angle.

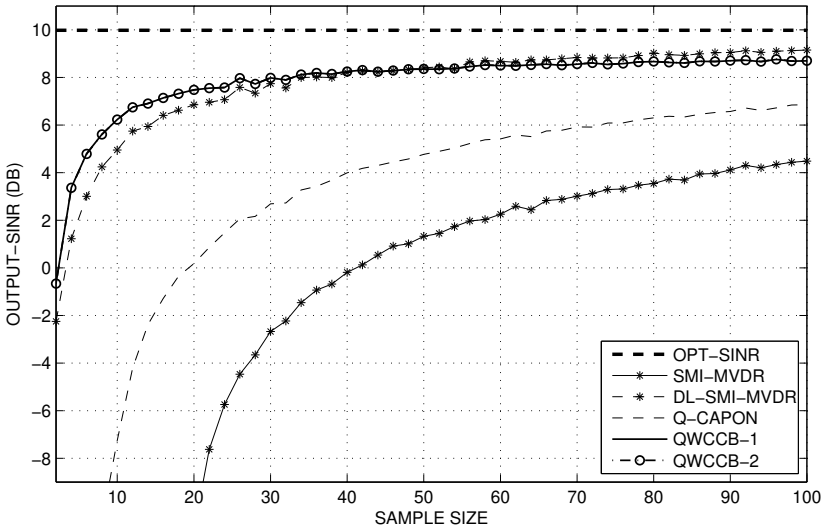

Figure C.2: Output-SINR versus sample size without steering vector mismatches: SNR=0 dB; the desired signal is from 5° with polarizations $(\gamma_{\rm d}, \eta_{\rm d}) = (15°, 50°)$, and interferences' polarizations are $(\gamma_{{\rm i},1}, \eta_{{\rm i},1}) = (35°, 80°)$ and $(\gamma_{{\rm i},2}, \eta_{{\rm i},2}) = (85°, 110°)$.

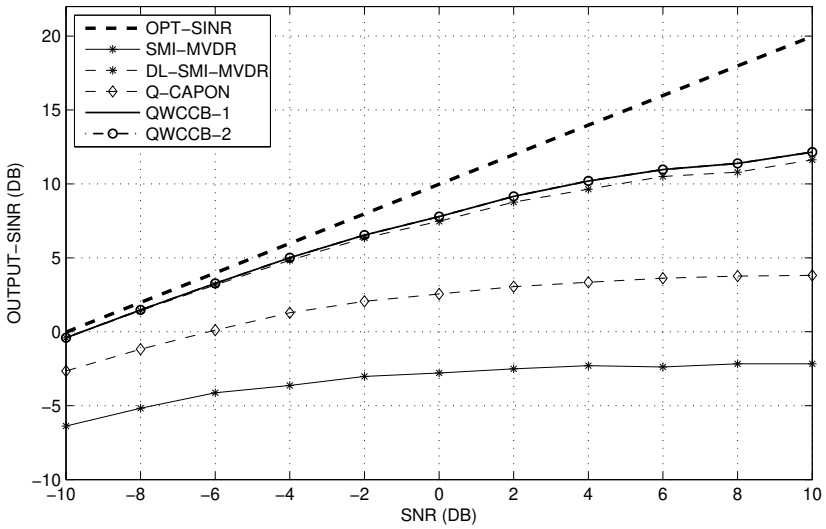

Figure C.3: Output-SINR versus SNR without steering vector mismatches: K=30; the desired signal is from 5° with polarizations $(\gamma_{\rm d}, \eta_{\rm d}) = (15°, 50°)$, and interferences' polarizations are $(\gamma_{\rm i,1}, \eta_{\rm i,1}) = (35°, 80°)$ and $(\gamma_{\rm i,2}, \eta_{\rm i,2}) = (85°, 110°)$.

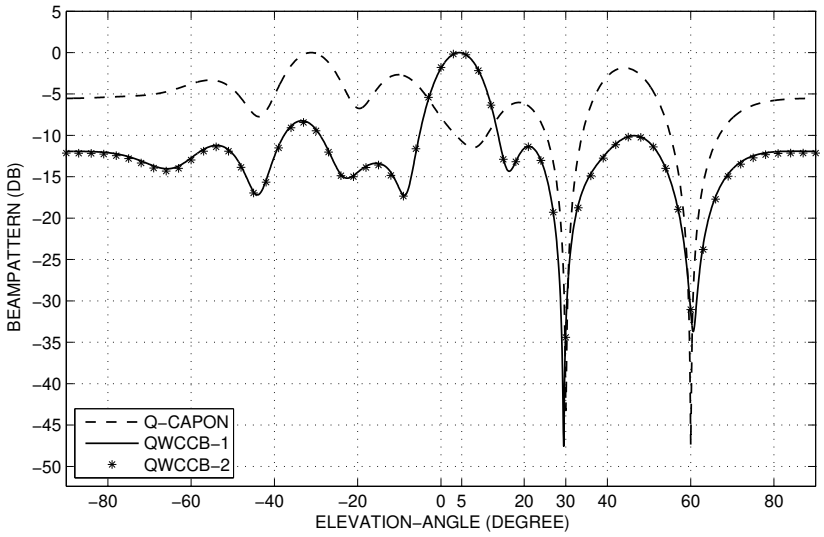

Figure C.4: Beam patterns of the proposed and the Q-Capon beamformers in the case of steering vector mismatches: SNR=0 dB, K=100; the desired signal is from 5° and two interferences are from 30° and 60°.

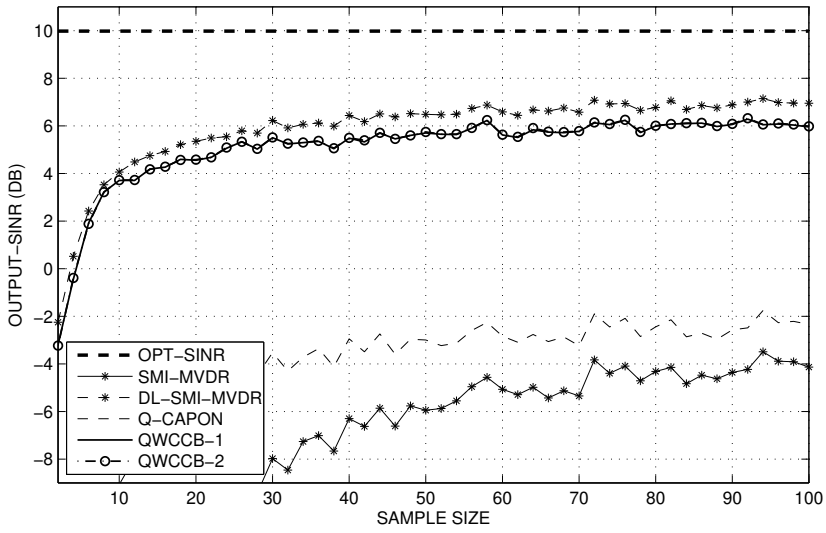

Figure C.5: Output-SINR versus sample size with steering vector mismatches: SNR=0 dB; the desired signal is from 5° with polarizations $(\gamma_{\rm d}, \eta_{\rm d}) = (15°, 50°)$, and interferences' polarizations are $(\gamma_{{\rm i},1}, \eta_{{\rm i},1}) = (35°, 80°)$ and $(\gamma_{{\rm i},2}, \eta_{{\rm i},2}) = (85°, 110°)$.

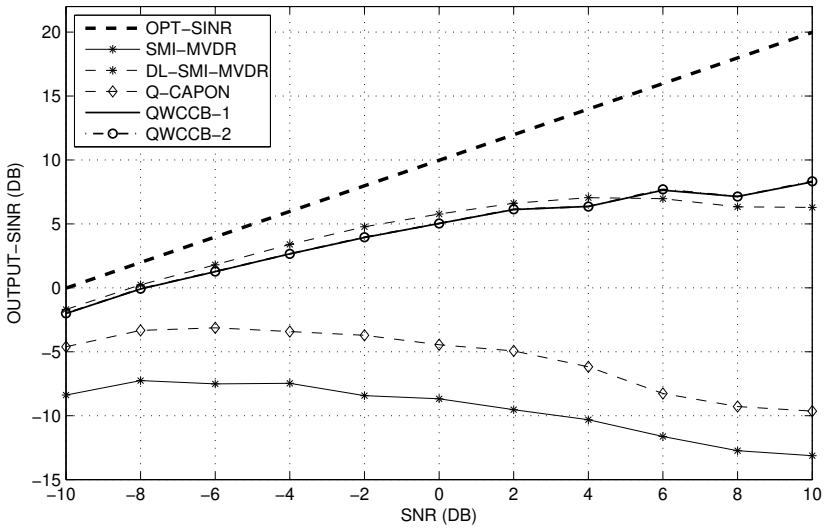

Figure C.6: Output-SINR versus SNR with steering vector mismatches: K=30; the desired signal is from 5° with polarizations $(\gamma_{\rm d}, \eta_{\rm d}) = (15^\circ, 50^\circ)$, and interferences' polarizations are $(\gamma_{\rm i,1}, \eta_{\rm i,1}) = (35^\circ, 80^\circ)$ and $(\gamma_{\rm i,2}, \eta_{\rm i,2}) = (85^\circ, 110^\circ)$.

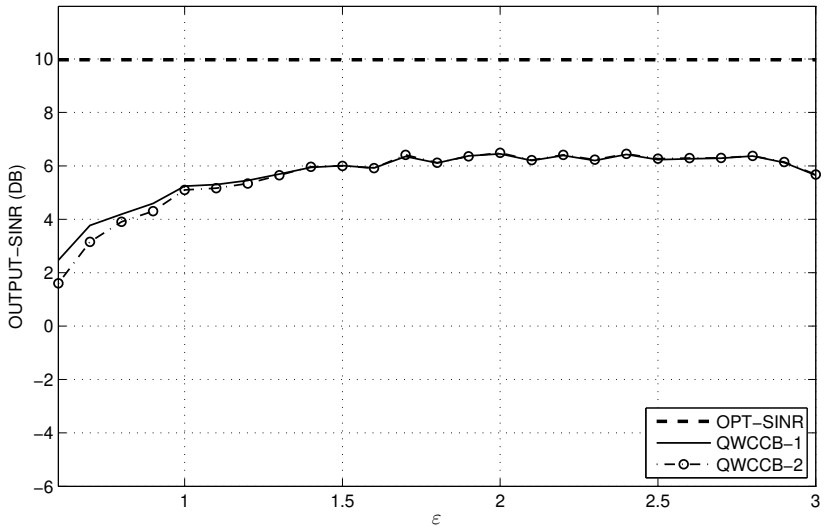

Figure C.7: Output-SINR versus ε with steering vector mismatches: SNR=0 dB, K=50; the desired signal is from 5° with polarizations $(\gamma_{\rm d}, \eta_{\rm d}) = (15°, 50°)$, and interferences' polarizations are $(\gamma_{{\rm i},1}, \eta_{{\rm i},1}) = (35°, 80°)$ and $(\gamma_{{\rm i},2}, \eta_{{\rm i},2}) = (85°, 110°)$.