On Extending the ADMM Algorithm to the Quaternion Algebra Setting

Srđan Lazendić*, Hendrik De Bie*, and Aleksandra Pižurica‡
*Department of Electronics and Information Systems, Clifford Research Group,
‡Department of Telecommunications and Information Processing, TELIN-GAIM
Faculty of Engineering and Architecture, Ghent University, Belgium
{Srdan.Lazendic; Hendrik.DeBie; Aleksandra.Pizurica}@UGent.be

***Abstract*—Many image and signal processing problems benefit from quaternion based models, due to their property of processing different features simultaneously. Recently the quaternion algebra model has been combined with the dictionary learning and sparse representation models. This led to solving versatile optimization problems over the quaternion algebra. Since the quaternions form a noncommutative algebra, calculation of the gradient of the quaternion objective function is usually fairly complex. This paper aims to present a generalization of the augmented directional method of multipliers over the quaternion algebra, while employing the results from the recently introduced generalized HR (GHR) calculus. Furthermore, we consider the convex optimization problems of real functions of quaternion variable.**

***Index Terms*—Numerical optimization, ADMM, Quaternions**

I. Introduction

The quaternion algebra was introduced by Hamilton in his seminal paper in 1843. Since then, it found applications in many different fields ranging from image processing [1], [2] and computer graphics [3] to design of space-time polarization block codes [4]. Since three imaginary units are well suited for representation of three color channels (R, G and B), recently the quaternions have been combined with dictionary learning methods [5], [6]. In all those applications there is a great interest for efficient numerical optimization algorithms.

In most of the current approaches to quaternion analysis, the main difference is in the definition of quaternion analyticity. The quaternion derivative is then defined for analytic quaternionic functions of quaternion varables [7]–[9]. In those settings one is limited only to a certain class of functions, in particular quaternion analytic functions. However, most of the functions in quaternion signal processing do not obey this property. Many of the optimization functions that appear in quaternion signal processing are non-analytic real cost functions of a quaternion variable. This problem was addressed in [10], where the authors connect the ℓ_1 and ℓ_2 minimization problems over the quaternion algebra with $L_{2,1}$ and ℓ_2 norm of the corresponding real matrix problems, respectively. However, this approach is applicable only when a suitable isomorphisms exist that faithfully represent the original problem in the real setting. For some important problems, such as the ℓ_0-norm minimization problem, finding the corresponding real representation in this setting is not evident. Thus, it is not possible to simply rewrite every quaternion function and operator as an equivalent real matrix problem.

A recently introduced new notion of the quaternion analyticity, also known as the generalized HR (GHR) calculus [11], allows working directly with real cost functions of quaternion variable. This approach represents the natural extension of the well-known complex Wirtinger calculus, which has been widely used in complex-valued signal processing. The GHR theory proves to be useful for versatile optimization problems which involve quaternion valued elements [12].

Certain optimization tools have been extended to the quaternion algebra in [13], in the context of audio separation. In particular, the authors there solve the Principal Component Pursuit (PCP) by using proximal operator that they define in the complex and quaternion domains. This approach is confined to the ℓ_1- and trace-norm regularization functions. Promixal operators for more complex optimization functions over quaternions have not been investigated yet.

Different algorithms that involve linearization or splitting of functions and/or variables are a special case of the proximal operator [14]. Among these is the alternating directional method of multipliers (ADMM) [15], [16] as a variant of the augmented Lagrangian scheme and has been widely used in signal/image processing, optimal control and machine learning. To the best of our knowledge ADMM has not been studied beyond real and complex numbers and thus is not in the scope of the quaternion algebra yet.

In this work, we propose the ADMM method over the quaternion algebra, which we call Q-ADMM. In particular, we show that our model naturally generalizes the classical ADMM algorithm. In the ADMM method, where the objective terms are being handled separately, the functions are accessed only through their proximal operators. Thus, there is a need for an elegant and efficient way of calculating derivatives of the objective quaternion function. We will show and explain that by using the GHR calculus, iterations in the proposed Q-ADMM algorithm can be obtained in a compact and intuitive way directly over the quaternion algebra. Finally, we illustrate on a couple of examples that the introduced algorithm quickly solves the optimization problems frequently appearing in quaternion optimization, signal processing, machine learning and statistics.

The organization of the paper is as follows. In Section II we give the preliminaries about quaternion linear algebra and GHR calculus. Section III contains the detailed explanation of the proposed Q-ADMM method. In the same section we show that the proposed method directly solves common examples from quaternion signal and image processing and illustrate it on an example of randomly generated data. Section IV concludes the paper.

II. Preliminaries

A. Quaternion algebra

The quaternion algebra is a 4-dimensional unital, distributive algebra over $\mathbb{R}$ with basis $\{1, \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ where 1 is the multiplicative identity and $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are imaginary units that satisfy $\mathbf{e}_i^2 = -1$. It holds that $\mathbf{e}_1\mathbf{e}_2 = \mathbf{e}_3$ and $\mathbf{e}_i\mathbf{e}_j = -\mathbf{e}_j\mathbf{e}_i$, for $i \neq j$, $i, j \in \{1, 2, 3\}$.

If $x \in \mathbb{H}$ is a quaternion element then it can be written as the sum of the real and imaginary part, i.e., as

$$x = x_0 + x_1\mathbf{e}_1 + \cdots + x_3\mathbf{e}_3 = x_0 + \sum_{i=1}^{3} x_i\mathbf{e}_i$$

where each $x_i \in \mathbb{R}$. The conjugate of a quaternion is defined as $\overline{x} = x_0 - \sum_{i=0}^{3} x_i\mathbf{e}_i$ and the norm is introduced as

$$|x| = \sqrt{x\overline{x}} = \sqrt{\overline{x}x} = \sqrt{x_0^2 + x_1^2 + x_2^2 + x_3^2}. \tag{1}$$

It is important to stress that the quaternion algebra is non-commutative but associative. This means that for arbitrary $x, y, z \in \mathbb{H}$: $\quad xy \neq yx$ but $x(yz) = (xy)z$. More about quaternion linear algebra can be found in [17]. It will also be useful to use the fact that for the real and complex part of x it holds that

$$2\text{Re}(x) = x + \overline{x} \quad \text{and} \quad 2\text{Im}(x) = x - \overline{x}. \tag{2}$$

B. Quaternion vectors and matrices

The quaternion vector $\mathbf{x} \in \mathbb{H}^{n\times 1}$ is a vector where each entry is a quaternion. Often it will be useful to write it as

$$\mathbf{x} = \mathbf{x}_0 + \mathbf{x}_1\mathbf{e}_1 + \mathbf{x}_2\mathbf{e}_3 + \mathbf{x}_3\mathbf{e}_3 \tag{3}$$

where now each vector $\mathbf{x}_i \in \mathbb{R}^{n\times 1}$ for $i = 1, \ldots, n$. Similarly, the quaternion matrix $\mathbf{A} = [a_{ij}]_{i,j=1}^{m,n} \in \mathbb{H}^{m\times n}$ is a matrix with entries in $\mathbb{H}$. Also, we can write it as

$$\mathbf{A} = \mathbf{A}_0 + \mathbf{A}_1\mathbf{e}_1 + \mathbf{A}_2\mathbf{e}_2 + \mathbf{A}_3\mathbf{e}_3, \tag{4}$$

where each $\mathbf{A}_i \in \mathbb{R}^{m\times n}$, for $i = 0, 1, 2, 3$.

The conjugation can be extended to vectors and matrices. The conjugate transpose matrix of a matrix $\mathbf{A} = [a_{ij}]_{i,j=1}^{m,n} \in \mathbb{H}^{m\times n}$ is defined as

$$\mathbf{A}^H = \overline{\mathbf{A}^T} = [\overline{a_{ji}}]_{j,i=1}^{n,m} \in \mathbb{H}^{n\times m}.$$

The inner product can be introduced as a function $\langle\cdot,\cdot\rangle : \mathbb{H}^n \times \mathbb{H}^n \to \mathbb{H}$ given by $\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}^H\mathbf{y} = \sum_{i=1}^{n} \overline{x_i}y_i$. Then, for a quaternion vector $\mathbf{x} \in \mathbb{H}^{n\times 1}$ we can define its norm by $\|\mathbf{x}\|_2^2 = \mathbf{x}^H\mathbf{x}$.

If $x = x_0 + x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \in \mathbb{H}$ is a quaternion, then we define the real-vector representation $\nu(x) \in \mathbb{R}^{4\times 1}$ as $\nu(x) = [x_0\ x_1\ x_2\ x_3]^T$. In this way the real-linear isomorphism $\nu : \mathbb{H} \to \mathbb{R}^{4\times 1}$ is obtained. By multiplying arbitrary $x, y \in \mathbb{H}$, there holds that

$$\nu(xy) = \chi(x)\nu(y) \tag{5}$$

where $\chi(x) \in \mathbb{R}^{4\times 4}$ is given by

$$\chi(x) = \begin{bmatrix} x_0 & -x_1 & -x_2 & -x_3 \\ x_1 & x_0 & -x_3 & x_2 \\ x_2 & x_3 & x_0 & -x_1 \\ x_3 & -x_2 & x_1 & x_0 \end{bmatrix}. \tag{6}$$

In this way we define the real, linear, injective map $\chi : \mathbb{H} \to \mathbb{R}^{4\times 4}$. The same maps can be extended to quaternion vectors and matrices and if we use the same notation this means that we have:

$$\nu : \mathbb{H}^{n\times 1} \to \mathbb{R}^{4n\times 1}, \quad \nu\left(\begin{bmatrix} x_0 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} \nu(x_0) \\ \vdots \\ \nu(x_n) \end{bmatrix}, \tag{7}$$

for $x_1, \ldots, x_n \in \mathbb{H}$ and if $a_{i,j} \in \mathbb{H}$ are entries of the matrix

$$\chi : \mathbb{H}^{m\times n} \to \mathbb{R}^{4m\times 4n}, \quad \chi\left([a_{i,j}]_{i,j=1}^{m,n}\right) = [\chi(a_{i,j})]_{i,j=1}^{m,n}. \tag{8}$$

Thus, we are able to define multiplication between quaternion vectors and quaternion matrices by using the real matrix multiplication. Indeed, for every $\mathbf{A} \in \mathbb{H}^{m\times n}$ and every $\mathbf{x} \in \mathbb{H}^{n\times 1}$ it holds that $\nu(\mathbf{A}\mathbf{x}) = \chi(\mathbf{A})\nu(\mathbf{x})$. Note that for $\mathbf{x} \in \mathbb{H}^{n\times 1}$ it holds $\|\mathbf{x}\|_2 = \|\nu(\mathbf{x})\|_2$. The proofs of all the matrix representations presented here can be found in [17].

C. Generalized HR (GHR) calculus

Most of the recent approaches to quaternion analysis are based on the HR calculus and its generalization GHR calculus, which can handle non-analytic functions directly in the quaternion domain [12]. Similar to the Wirtinger complex calculus [18], in the GHR calculus the derivatives are being taken with respect to a quaternion variable and its involutions. The main advantage of this approach is that the derivatives of quaternion matrix functions can be calculated directly in an elegant way, without using the isomorphism with real vectors. For our work here, the derivatives of the real function of quaternion variables, summarized in Table I are relevant. Furthermore, the GHR derivatives allow us to find stationary points for scalar real valued functions of a quaternion variable. It was shown in [12], that the necessary condition for a solution in nonlinear optimization to be optimal is

$$\mathcal{D}_{\mathbf{x}}f(\mathbf{x}) = \mathcal{D}_{\overline{\mathbf{x}}}f(\mathbf{x}) = 0, \tag{9}$$

where $\mathcal{D}$ is the left matrix derivative and $f : \mathbb{H}^{n\times 1} \to \mathbb{R}$.

III. Proposed model and results

A. Alternating Directional Method of Multipliers over $\mathbb{H}$

We shall now extend the ADMM algorithm such that it can be directly applied to the optimization problems over the quaternion algebra. This method will be called Q-ADMM. Here we will motivate and show how some of the most

TABLE I
MATRIX DERIVATIVES OF FUNCTIONS OF TYPE $f(\mathbf{x})$

$f(\mathbf{x})$	$\mathcal{D}_{\mathbf{x}} f$	Remark
$\mathbf{x}^H\mathbf{A}\mathbf{x}$	$\mathbf{x}^H\mathbf{A} - \frac{1}{2}(\mathbf{A}\mathbf{x})^H$	$\mathbf{A} \in \mathbb{H}^{n\times n}$, $\mathbf{x} \in \mathbb{H}^{n\times 1}$
$\mathbf{x}^H\mathbf{A}$	$-\frac{1}{2}\mathbf{A}^H$	$\mathbf{A} \in \mathbb{H}^{n\times n}$, $\mathbf{x} \in \mathbb{H}^{n\times 1}$
$\mathbf{A}\mathbf{x}$	$\mathbf{A}$	$\mathbf{A} \in \mathbb{H}^{n\times n}$, $\mathbf{x} \in \mathbb{H}^{n\times 1}$
$\mathbf{A}\overline{\mathbf{x}}$	$-\frac{1}{2}\mathbf{A}$	$\mathbf{A} \in \mathbb{H}^{n\times n}$, $\mathbf{x} \in \mathbb{H}^{n\times 1}$
$\mathbf{x}^H\mathbf{b}$	$-\frac{1}{2}\mathbf{b}^H$	$\mathbf{b} \in \mathbb{H}^{n\times 1}$, $\mathbf{x} \in \mathbb{H}^{n\times 1}$

common optimization problems over the quaternion algebra can be seen in this framework.

We assume as in the classical case, that the Q-ADMM method over $\mathbb{H}$ solves problems of the form

$$\begin{aligned} \min_{\mathbf{x},\mathbf{z}} \quad & f(\mathbf{x}) + g(\mathbf{z}) \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z} = \mathbf{c} \end{aligned} \tag{10}$$

with variables $\mathbf{x} \in \mathbb{H}^{n\times 1}$, $\mathbf{z} \in \mathbb{H}^{m\times 1}$, $\mathbf{c} \in \mathbb{H}^{p\times 1}$, $\mathbf{A} \in \mathbb{H}^{p\times n}$ and $\mathbf{B} \in \mathbb{H}^{p\times m}$, where we assume that $f : \mathbb{H}^{n\times 1} \to \mathbb{R}$ and $g : \mathbb{H}^{m\times 1} \to \mathbb{R}$. In this article we restrict ourselves only to the case when the functions f and g are convex.

By directly following the real case approach, we should form the Lagrangian function. But this cannot be done directly, because the constraint $\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z} = \mathbf{c}$ is quaternion valued. Since the quaternions are not an ordered field, we cannot directly apply the same idea, but we can use the following trick which makes the Lagrangian function real valued:

$$\mathcal{L}_0(\mathbf{x},\mathbf{z},\mathbf{y}) = f(\mathbf{x}) + g(\mathbf{z}) + 2\mathrm{Re}\left(\mathbf{y}^H(\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z} - \mathbf{c})\right). \tag{11}$$

The augmented Lagrangian function for the problem (10) can then be defined by

$$\mathcal{L}_\rho(\mathbf{x},\mathbf{z},\mathbf{y}) = \mathcal{L}_0(\mathbf{x},\mathbf{z},\mathbf{y}) + \frac{\rho}{2}\|\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z} - \mathbf{c}\|_2^2, \tag{12}$$

where $\rho > 0$ is the penalty parameter and $\mathbf{y} \in \mathbb{H}^{p\times 1}$ is the Lagrange multiplier. The ADMM algorithm consists of the following steps

$$\mathbf{x}^{k+1} = \arg\min_{\mathbf{x}} \mathcal{L}_\rho(\mathbf{x}, \mathbf{z}^k, \mathbf{y}^k), \tag{13}$$

$$\mathbf{z}^{k+1} = \arg\min_{\mathbf{z}} \mathcal{L}_\rho(\mathbf{x}^{k+1}, \mathbf{z}, \mathbf{y}^k), \tag{14}$$

$$\mathbf{y}^{k+1} = \mathbf{y}^k + \rho(\mathbf{A}\mathbf{x}^{k+1} + \mathbf{B}\mathbf{z}^{k+1} - \mathbf{c}). \tag{15}$$

Let us define the residual $\mathbf{r} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{z} - \mathbf{c}$. The Lagrangian function contains the term $2\mathrm{Re}(\mathbf{r})$, which we will first rewrite by using (2). We have that $2\mathrm{Re}\left(\mathbf{y}^H\mathbf{r}\right) = \mathbf{y}^H\mathbf{r} + \mathbf{r}^H\mathbf{y} \in \mathbb{R}$, which is suitable for application of the GHR differentiation.

B. Examples and simulations

In the sequel, we show that the introduced method is well-defined and that the introduced theory has a practical applications. Thus, let us consider a few important examples in order to show how we can solve them by using the Q-ADMM method.

Example 1: The quaternion collaborative representation-based classification (QCRC) model is intensively used for the color image classification [10]. Given a test sample $\mathbf{b}$ and a quaternion dictionary $\mathbf{A}$, we want to obtain the quaternion representation vector $\mathbf{x}$. The corresponding minimization problem is given by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_2^2.$$

The solution to this problem in [10] is based on the existing isomorphism between quaternion and real vectors. This approach leads to tedious calculations and to implementations of huge real matrices. Now we will solve the same problem in a more elegant way by the proposed ADMM method over $\mathbb{H}$. By introducing the variable $\mathbf{z} = \mathbf{b} - \mathbf{A}\mathbf{x}$ we obtain the problem in the form (10)

$$\begin{aligned} \min_{\mathbf{x},\mathbf{z}} \quad & \lambda\|\mathbf{x}\|_2^2 + \|\mathbf{z}\|_2^2 \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} + \mathbf{z} = \mathbf{b}. \end{aligned}$$

The augmented Lagrangian is then given by

$$\begin{aligned} \mathcal{L}_\rho(\mathbf{x},\mathbf{z},\mathbf{y}) &= \lambda\|\mathbf{x}\|_2^2 + \|\mathbf{z}\|_2^2 + \mathbf{y}^H\left(\mathbf{A}\mathbf{x} + \mathbf{z} - \mathbf{b}\right) \\ &+ \left(\mathbf{x}^H\mathbf{A}^H + \mathbf{z}^H - \mathbf{b}^H\right)\mathbf{y} + \frac{\rho}{2}\|\mathbf{A}\mathbf{x} + \mathbf{z} - \mathbf{b}\|_2^2 \\ &= \lambda\mathbf{x}^H\mathbf{x} + \mathbf{z}^H\mathbf{z} + \mathbf{y}^H(\mathbf{A}\mathbf{x} + \mathbf{z} - \mathbf{b}) \\ &+ \left(\mathbf{x}^H\mathbf{A}^H + \mathbf{z}^H - \mathbf{b}^H\right)\mathbf{y} \\ &+ \frac{\rho}{2}(\mathbf{A}\mathbf{x} + \mathbf{z} - \mathbf{b})^H(\mathbf{A}\mathbf{x} + \mathbf{z} - \mathbf{b}). \end{aligned}$$

Now we can minimize each variable separately and conduct the ADMM iterations as follows. In order to obtain the updates for $\mathbf{x}$, $\mathbf{z}$ and $\mathbf{y}$ we calculate the GHR derivatives of the objective function by using the rules from Table I. By equating those with zero because of (9), we obtain the following ADMM updates:

$$\begin{aligned} \mathbf{x}^{k+1} &= (2\lambda\mathbb{I} + \rho\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H(\rho(\mathbf{c} - \mathbf{z}^k) - 2\mathbf{y}^k), \\ \mathbf{z}^{k+1} &= \frac{1}{2+\rho}(\rho(\mathbf{c} - \mathbf{A}\mathbf{x}^k) - 2\mathbf{y}^k), \\ \mathbf{y}^{k+1} &= \mathbf{y}^k + \rho(\mathbf{A}\mathbf{x}^k + \mathbf{z}^k - \mathbf{b}). \end{aligned}$$

Note that these updates are similar to those in the classical real case, which is not surprising given that this model represents its generalization.

Let us now consider a numerical simulation with random data where we directly work with the real-valued functions of the quaternion variable. We generate a random dictionary (overcomplete basis set) $\mathbf{A} \in \mathbb{H}^{64\times 256}$ and different noisy and clear signals $\mathbf{b} \in \mathbb{H}^{64\times 1}$. The obtained simulation results demonstrate that the ADMM can quickly solve the considered optimization problem and obtain satisfying results after only a few dozen iterations. Detailed mathematical proofs of the residual, objective and dual variable convergence results can be obtained and will follow in our future work.

The regularization parameters λ and ρ, can be optimized by grid search but that is not our focus here. The value of the objective function as well as the residual and dual-variable

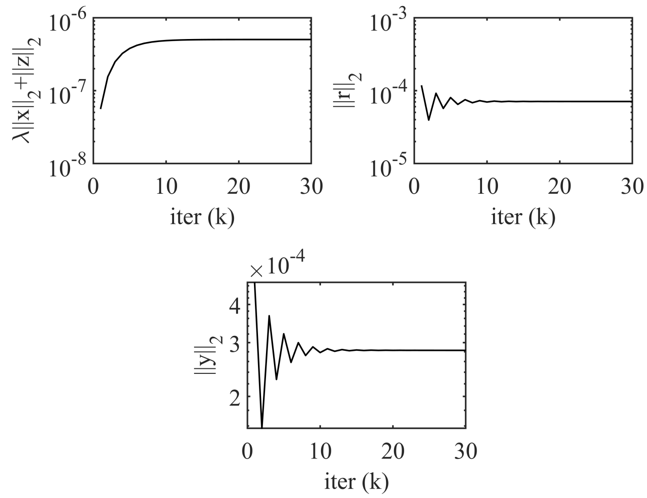

Fig. 1. The objective value, residual and dual variable convergence per iteration k, for $\lambda = 10^{-2}$ and $\rho = 3$.

values for 30 iterations, for $\lambda = 10^{-2}$ and $\rho = 3$ are plotted in Fig. 1.

Example 2: Another important convex example that appears in practice is the ℓ_1-norm minimization problem. The sparse representation classification (SRC) in the quaternion setting was used in [10] for color face recognition. The quaternion Lasso (QLasso) model computes the quaternion sparse representation vector by solving

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1. \tag{16}$$

We will now see that the advantage of the introduced Q-ADMM over $\mathbb{H}$ is even clearer, since the real transformed method [10] is given by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbb{H}^{n\times 1}} \|\nu(\mathbf{b}) - \chi(\mathbf{A})\nu(\mathbf{x})\|_2 + \lambda\|\mathcal{R}(\mathbf{x})\|_{2,1}, \tag{17}$$

and leads to the group Lasso model with huge real matrices. Problem (17) is then solved by using the classical real ADMM method. We formulate the problem (16) in the form suitable for the application of Q-ADMM by introducing a new variable $\mathbf{z}$. The problem then becomes

$$\begin{aligned} \min_{\mathbf{x},\mathbf{z}} \quad & \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2^2 + \lambda\|\mathbf{z}\|_1 \\ \text{s.t.} \quad & \mathbf{x} - \mathbf{z} = 0. \end{aligned}$$

In this form the problem can be directly solved by using Q-ADMM since it is represented in the form where the GHR calculus can be applied and the iterations of the Q-ADMM can easily be obtained. By using the differentiation operators from [11] and [12] we are able to efficiently calculate the derivatives of the $\|\cdot\|_1$-function.

Based on the presented examples, we can observe that the proposed Q-ADMM method can be used for solving the optimization problems over the quaternion algebra by directly working with the functions of a quaternion variable. This makes it suitable for numerous optimization problems which involve the quaternion valued parameters. This way we are able to solve optimization problems which cannot simply be rewritten as their equivalent real optimization problems. In the future, we will further explore the potential of the Q-ADMM.

IV. Conclusion

In this paper, we extend the well-known ADMM algorithm to the quaternion setting, which can then be applied to many optimization problems over the quaternion algebra. The main ideas behind the Q-ADMM model were presented. We also stressed the importance of being able to solve those problems directly without using the isomorphism with the algebra of real numbers. The presented examples together with the experimental results show that the proposed algorithm together with the GHR calculus, has potential to be used for different quaternion minimization problems.

Acknowledgment

This research has been partially supported by the institutional grant no. BOF15/24J/078 of Ghent University.

References

[1] T. A. Ell, N. Le Bihan, and S. J. Sangwine, *Quaternion Fourier transforms for signal and image processing*. John Wiley & Sons, 2014.

[2] Z. Jia, M. Ng, and G. Song, "A quaternion framework for robust color images completion," *Preprint, May*, 2018.

[3] R. Mukundan, "Quaternions: From classical mechanics to computer graphics, and beyond," in *Proceedings of the 7th Asian Technology conference in Mathematics*, 2002, pp. 97–105.

[4] H. Ma, H. Kan, and H. Imai, "A new quaternion design for space-time-polarization block code with full diversity," *IEICE transactions on communications*, vol. 92, no. 2, pp. 671–674, 2009.

[5] L. Yu, Y. Xu, H. Xu, and H. Zhang, "Quaternion-based sparse representation of color image," in *2013 International Conference on Multimedia and Expo (ICME)*. IEEE, 2013, pp. 1–7.

[6] Y. Xu, L. Yu, H. Xu, H. Zhang, and T. Nguyen, "Vector sparse representation of color image using quaternion matrix analysis," *IEEE Transactions on image processing*, vol. 24, no. 4, pp. 1315–1329, 2015.

[7] A. Sudbery, "Quaternionic analysis," in *Math. Proc. Camb. Phil. Soc*, vol. 85, no. 2, 1979, pp. 199–225.

[8] G. Gentili and D. C. Struppa, "A new theory of regular functions of a quaternionic variable," *Advances in Mathematics*, vol. 216, no. 1, pp. 279–301, 2007.

[9] F. Colombo, I. Sabadini, and D. C. Struppa, *Entire slice regular functions*. Springer, 2016.

[10] C. Zou, K. I. Kou, and Y. Wang, "Quaternion collaborative and sparse representation with application to color face recognition," *IEEE Transactions on image processing*, vol. 25, no. 7, pp. 3287–3302, 2016.

[11] D. Xu, C. Jahanchahi, C. C. Took, and D. P. Mandic, "Enabling quaternion derivatives: the generalized hr calculus," *Royal Society open science*, vol. 2, no. 8, p. 150255, 2015.

[12] D. Xu and D. P. Mandic, "The theory of quaternion matrix derivatives," *IEEE Transactions on Signal Processing*, vol. 63, no. 6, pp. 1543–1556, 2015.

[13] T.-S. T. Chan and Y.-H. Yang, "Complex and quaternionic principal component pursuit and its application to audio separation," *IEEE Signal Processing Letters*, vol. 23, no. 2, pp. 287–291, 2016.

[14] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-point algorithms for inverse problems in science and engineering*. Springer, 2011, pp. 185–212.

[15] M. R. Hestenes, "Multiplier and gradient methods," *Journal of optimization theory and applications*, vol. 4, no. 5, pp. 303–320, 1969.

[16] J. Eckstein and D. P. Bertsekas, "On the douglas—rachford splitting method and the proximal point algorithm for maximal monotone operators," *Mathematical Programming*, vol. 55, no. 1-3, pp. 293–318, 1992.

[17] L. Rodman, *Topics in quaternion linear algebra*. Princeton University Press, 2014, vol. 45.

[18] W. Wirtinger, "Zur formalen theorie der funktionen von mehr komplexen veränderlichen," *Mathematische Annalen*, vol. 97, no. 1, pp. 357–375, 1927.