Quaternion Basis Pursuit via Iteratively Reweighted Least Squares

XUE-YAN TIAN[1], CUI-MING ZOU[1]

[1]College of Science, Huazhong Agricultural University, Wuhan, 430070, China
E-MAIL: tianxy@webmail.hzau.edu.cn, zoucuiming2006@163.com

Abstract:

Sparse signal recovery is a crucial topic in signal processing. While ℓ_1 regularization methods are widely used, they face significant computational challenges when dealing with high-dimensional nonsmooth functions. To address this issue, this paper proposes a Quaternion Basis Pursuit (QBP) method, which achieves sparse decomposition by minimizing the ℓ_1 norm of quaternion signals. We propose an efficient optimization strategy based on Iteratively Reweighted Least Squares (IRLS) and validate its effectiveness through extensive experiments on quaternion signal recovery tasks. The experimental results demonstrate that QBP achieves better performance for quaternion signal reconstruction while significantly reducing the recovery errors. It provides a novel way for sparse optimization in the quaternion domain, offering broad application potential.

Keywords:

Sparse Signal; Iteratively Reweighted Least Squares; Quaternion Basis Pursuit

1 Introduction

Sparse signal recovery has garnered significant attention in recent years since its wide applications in signal processing, image reconstruction, compressed sensing, and machine learning. The task of sparse signal recovery is to recover the original sparse signal from the known measurement matrix and observation vector [1]. To obtain a sparse signal $\mathbf{x}$, the l_0 minimization constraint is usually used to solve a problem as:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}, \tag{1}$$

which is NP-hard in general [2]. The computational complexity will increase rapidly with the increase of sparse vector dimension. For computational efficiency, it is common to replace ℓ_0 minimization with its convex relaxation as:

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{2}$$

Here, $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{y} \in \mathbb{R}^m$ are given. It is commonly referred to as ℓ_1 norm minimization [3] or basis pursuit [4] in the literature.

Basis pursuit (BP) transforms the non-convex ℓ_0 norm optimization into a convex problem, facilitating global optimal approximation. Theory guarantees have been established for BP, particularly under the Restricted Isometry Property (RIP), which ensures the solution obtained from the convex relaxation remains equivalent to the original ℓ_0 minimization problem [5]. However, despite these theoretical guarantees, BP may still exhibit limitations in practical scenarios, such as suboptimal recovery in highly sparse settings. To address this issue, researchers have proposed various improved methods. Among them, the Iterative Reweighted Least Squares (IRLS) algorithm is an important enhancement technique. By introducing adaptive weights and solving iteratively, IRLS aims to better approximate the ℓ_0 norm and improve the accuracy and recovery performance of sparse solutions.

The IRLS method is a widely used numerical optimization technique, specially in sparse signal recovery and compressed sensing [6]. By dynamically updating weights at each iteration, IRLS transforms the original non-convex optimization problem into a series of weighted least squares problems, making it easier to solve. In particular, IRLS has been extensively applied in sparse signal recovery and compressed sensing , where it is used to minimize the ℓ_p $(0 < p < 1)$ norm [7] and achieve sparser solutions than ℓ_1 methods [6]. Moreover, IRLS is useful in robust regression, where it assigns different weights to residuals to effectively reduce the influence of outliers [2]. Previous BP studies have shown that under certain conditions, IRLS enjoys a linear convergence rate and efficiently approximates the optimal solution [8].

The outlines of the paper is organized as follows. Introduce basic knowledge on quaternion and basis pursuit in Section 2. In section 3, we have described model and its optimization process. Section 4 presents the experimental results. Finally, conclude the paper in section 5.

2 Related Works

Now, we introduce the basics related to quaternion and BP.

2.1 Quaternion

A quaternion $\dot{q} \in \mathbb{H}$ is a type of hyper-complex number that consists of a real scalar part and three imaginary components [9]. It is represented as $\dot{q} = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$, where $q_0, q_1, q_2, q_3 \in \mathbb{R}$. $\mathrm{i}, \mathrm{j}, \mathrm{k}$ are imaginary units that satisfy the multiplication rules: $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1$. This property leads to the non-commutative nature of quaternion multiplication, meaning that for any two quaternions $\dot{p}$ and $\dot{q}$, $\dot{p}\dot{q} \neq \dot{q}\dot{p}$. The conjugate transport of a quaternion $\dot{q}$ is defined as $\bar{\dot{q}} = q_0 - q_1\mathrm{i} - q_2\mathrm{j} - q_3\mathrm{k}$. The modulus of a quaternion $\dot{q}$ is given by $|\dot{q}| = \sqrt{\bar{\dot{q}}\dot{q}}$. A quaternion vector $\dot{\mathbf{q}}$ can be expressed in the form $\dot{\mathbf{q}} = \mathbf{q}_0 + \mathbf{q}_1\mathrm{i} + \mathbf{q}_2\mathrm{j} + \mathbf{q}_3\mathrm{k}$, where each component $\mathbf{q}_l \in \mathbb{R}^n$ for $l = 0, 1, 2, 3$. For a quaternion vector $\dot{\mathbf{q}} = [\dot{q}_1, \dot{q}_2, \ldots, \dot{q}_n]^T$, its conjugate transpose is defined as $\dot{\mathbf{q}}^H = [\bar{\dot{q}}_1, \bar{\dot{q}}_2, \ldots, \bar{\dot{q}}_n]$. The ℓ_1 norm of $\dot{q}$ is given by $\|\dot{\mathbf{q}}\|_1 = \sum_i |\dot{q}_i|$. For two quaternion vectors $\dot{\mathbf{p}}$ and $\dot{\mathbf{q}}$, their inner product is expressed as $\langle \dot{\mathbf{p}}, \dot{\mathbf{q}} \rangle = \dot{\mathbf{p}}^H\dot{\mathbf{q}}$. A quaternion matrix $\dot{\mathbf{Q}} = [\dot{\mathbf{q}}_1, \dot{\mathbf{q}}_2, \cdots, \dot{\mathbf{q}}_n] \in \mathbb{H}^{m\times n}$. The conjugate transport of $\dot{\mathbf{Q}}$ is given by $\dot{\mathbf{Q}}^H = [\dot{\mathbf{q}}_1^H, \dot{\mathbf{q}}_2^H, \cdots, \dot{\mathbf{q}}_n^H]$.

2.2 Basis Pursuit

Basis Pursuit (BP) is an optimization method used for sparse signal representation. It employs convex optimization to find the sparsest signal representation in an overcomplete dictionary [4]. The fundamental idea is to solve a convex ℓ_1 minimization problem, which promotes sparsity while ensuring a unique and stable solution. Building on ℓ_1 norm minimization, several variants have been proposed to enhance its performance in different scenarios. For instance, the L1-L2 [10] minimization method introduces an additional regularization term to balance sparsity and stability, enabling the accurate recovery of sparse vectors. The L1/L2 model [11], known for its scale-invariance property, improves adaptability. For multi-channel and noisy environments, QSR [12] provides a more robust solution, QEN [13] is particularly effective in handling signals with high inter-group correlation. Depending on prior knowledge about the signal, researchers have developed extensions of BP, such as block BP methods [14] for structured sparsity and robust BP [15] techniques for handling noise and outliers. Despite these advancements, BP has seen limited application in multi-channel settings, showing a research gap that needs more study.

3 Quaternion Basis Pursuit

Although the Basis Pursuit (BP) method is effective, in practical applications, such as color image processing, the data is often multi-channel. Traditional BP can only process each channel separately and then combine them, which results in the loss of inter-channel information. Therefore, there is a need to extend BP to multi-channel data. Previous literature has shown that quaternion-based methods offer advantages in handling such multi-channel data [16]. To address the above issue, a quaternion basis pursuit (QBP) model is proposed as follows

$$\min_{\dot{\mathbf{x}}\in\mathbb{H}^n} \|\dot{\mathbf{x}}\|_1 \quad \text{s.t.} \quad \dot{\mathbf{y}} = \dot{\mathbf{A}}\dot{\mathbf{x}}, \tag{3}$$

where $\|\dot{\mathbf{x}}\|_1$ denotes the quaternion ℓ_1 norm, $\dot{\mathbf{A}}$ is measurement matrix and $\dot{\mathbf{y}}^*$ is observation vector. Considering $\|\dot{\mathbf{x}}\|_1$ is non-smooth and non-convex, we extend the Majorization-Minimization (MM) algorithm [17] from the real domain to the quaternion domain. Following the ideas in the literature [18], a function $\mathcal{J}_\epsilon(\dot{\mathbf{x}})$ is introduced, which serves as a smooth approximation of ℓ_1 norm. For a fixed smoothing parameter $\epsilon > 0$, it is defined as $\mathcal{J}_\epsilon(\dot{\mathbf{x}}) := \sum_{i=1}^n J_\epsilon(\dot{x}_i)$ with

$$J_\epsilon(\dot{x}_i) := \begin{cases} |\dot{x}_i|, & |\dot{x}_i| > \epsilon, \\ \frac{1}{2}\left(\frac{\dot{x}_i^2}{\epsilon} + \epsilon\right), & |\dot{x}_i| < \epsilon, \end{cases} \tag{4}$$

where $\dot{\mathbf{x}} = [\dot{x}_1, \dot{x}_2 \cdots, \dot{x}_n]^T$. Since the modulus of each component of $\dot{\mathbf{x}}$ is smooth in the real domain, it remains smooth in the quaternion domain. However, the function is still non-convex. Due to its non-convexity, we approximate it using a suitable quadratic function $Q_\epsilon(\cdot, \dot{\mathbf{x}})$ defined by

$$\begin{aligned} Q_\epsilon(\dot{\mathbf{z}}, \dot{\mathbf{x}}) :=& \mathcal{J}_\epsilon(\dot{\mathbf{x}}) + \langle \nabla\mathcal{J}_\epsilon(\dot{\mathbf{x}}), \dot{\mathbf{z}} - \dot{\mathbf{x}} \rangle + \frac{1}{2}\langle (\dot{\mathbf{z}} - \dot{\mathbf{x}}), \mathbf{W}(\dot{\mathbf{z}} - \dot{\mathbf{x}}) \rangle \\ =& \mathcal{J}_\epsilon(\dot{\mathbf{x}}) + \frac{1}{2}\langle \dot{\mathbf{z}}, \mathbf{W}\dot{\mathbf{z}} \rangle - \frac{1}{2}\langle \dot{\mathbf{x}}, \mathbf{W}\dot{\mathbf{x}} \rangle, \end{aligned} \tag{5}$$

where $\mathrm{diag}(\mathbf{w}_\epsilon(\dot{\mathbf{x}}))$ is denoted as $\mathbf{W}$. The gradient of J_ϵ at $\dot{\mathbf{x}}$, $\nabla J_\epsilon(\dot{\mathbf{x}})$, is defined by

$$\nabla J_\epsilon(\dot{x}_i) = \begin{cases} \dot{x}_i/|\dot{x}_i|, & |\dot{x}_i| > \epsilon, \\ \dot{x}_i/\epsilon, & |\dot{x}_i| \le \epsilon, \end{cases} \tag{6}$$

and $w_\epsilon(\dot{\mathbf{x}}) \in \mathbb{R}^n$ is a weight vector.
The function $Q_\epsilon(\cdot, \cdot)$ has the following properties:
(i) $\mathrm{diag}\left(\mathbf{w}_\epsilon(\dot{\mathbf{x}})\right)\dot{\mathbf{x}} = \nabla\mathcal{J}_\epsilon\left(\dot{\mathbf{x}}\right)$.
(ii) $Q_\epsilon(\dot{\mathbf{z}}, \dot{\mathbf{x}}) \ge \mathcal{J}_\epsilon(\dot{\mathbf{z}})$, the equation holds if and only if $\dot{\mathbf{z}} = \dot{\mathbf{x}}$.
Using these properties, we can update $\dot{\mathbf{x}}$.

During the update process of $\dot{\mathbf{x}}$, it aims to minimize $Q_\epsilon(\cdot, \dot{\mathbf{x}})$, which requires to find $\dot{\mathbf{z}}$ that minimizes $\langle \dot{\mathbf{z}}, \mathrm{diag}(\mathbf{w}_\epsilon(\dot{\mathbf{x}})\dot{\mathbf{z}}\rangle$. The weighted structure imposed by $\mathbf{w}_\epsilon(\dot{\mathbf{x}})$ plays a crucial role in refining the solution, guiding $\dot{\mathbf{z}}$ towards a sparser representation. Consequently, the update rule for $\dot{\mathbf{x}}$ can be written as

$$\dot{\mathbf{x}}^{(k+1)} = \boldsymbol{Q}_{inv}\left(\boldsymbol{P}\left(\mathbf{W}^{(k)^{-1}}\dot{\mathbf{A}}^H\right)\left(\boldsymbol{P}\left(\dot{\mathbf{A}}\mathbf{W}^{(k)^{-1}}\dot{\mathbf{A}}^H\right)\right)^{-1}\boldsymbol{Q}(\dot{\mathbf{y}})\right). \quad (7)$$

Here, for a matrix $\dot{\mathbf{M}} = \mathbf{M_0} + \mathbf{M_1}\mathrm{i} + \mathbf{M_2}\mathrm{j} + \mathbf{M_3}\mathrm{k} \in \mathbb{H}^{m\times n}$, the operator $\boldsymbol{P}$ [12] is defined as

$$\boldsymbol{P}(\dot{\mathbf{M}}) := \begin{bmatrix} \mathbf{M_0} & -\mathbf{M_1} & -\mathbf{M_2} & -\mathbf{M_3} \\ \mathbf{M_1} & \mathbf{M_0} & -\mathbf{M_3} & \mathbf{M_2} \\ \mathbf{M_2} & \mathbf{M_3} & \mathbf{M_0} & -\mathbf{M_1} \\ \mathbf{M_3} & -\mathbf{M_2} & \mathbf{M_1} & \mathbf{M_0} \end{bmatrix}.$$

For a vector $\dot{\mathbf{v}} = \mathbf{v_0} + \mathbf{v_1}\mathrm{i} + \mathbf{v_2}\mathrm{j} + \mathbf{v_3}\mathrm{k} \in \mathbb{H}^n$, the operator $\boldsymbol{Q}$ [12] transforms $\dot{\mathbf{v}}$ into a vector in $\mathbb{R}^{4n}$, i.e., $\boldsymbol{Q}(\dot{\mathbf{v}}) := [\mathbf{v}_0^{\mathrm{T}}\ \mathbf{v}_1^{\mathrm{T}}\ \mathbf{v}_2^{\mathrm{T}}\ \mathbf{v}_3^{\mathrm{T}}]^{\mathrm{T}}$, and $\boldsymbol{Q}_{inv}$ is its inverse transformation. $\boldsymbol{P}$ and $\boldsymbol{Q}$ has the following proposition: $\boldsymbol{Q}(\dot{\mathbf{M}}\dot{\mathbf{v}}) = \boldsymbol{P}(\dot{\mathbf{M}})\boldsymbol{Q}(\dot{\mathbf{v}}),\ \forall\, \dot{\mathbf{M}} \in \mathbb{H}^{m\times n},\ \forall\, \dot{\mathbf{v}} \in \mathbb{H}^n$;

With the fixed $\dot{\mathbf{x}}$, the followings perform updates for ϵ and $\mathbf{W}$. There are two methods for updating the smoothing parameter ϵ [19]: one uses a fixed value, and the other updates ϵ dynamically. In our approach, a dynamic update rule for ϵ is used , where ϵ follows a non-increasing sequence. This dynamic update strategy allows ϵ to gradually decrease, leading to a better approximation of the original ℓ_1 norm while maintaining numerical stability in the early iterations. Specifically, the updating rule is given by

$$\epsilon^{(k+1)} := \min\left\{\epsilon_k, \frac{\sigma_s(\dot{\mathbf{x}}^{(k+1)})_{\ell_1}}{n}\right\}, \quad (8)$$

where $\sigma_s(\dot{\mathbf{x}})_{\ell_1} = \inf\left\{\|\dot{\mathbf{x}} - \dot{\mathbf{z}}\|_1 : \dot{\mathbf{z}} \in \mathbb{H}^n \text{ and } \dot{\mathbf{z}} \text{ is } s\text{-sparse}\right\}$.

For each component $\dot{x}_i$, the weight w_i depends on its module. When $|\dot{x}_i| > \epsilon$, w_i is set to $1/|\dot{x}_i|$; when $|\dot{x}_i| \le \epsilon$, w_i is set to $1/\epsilon$. The update rule is as follows

$$\begin{aligned} &\mathbf{W} = \mathrm{diag}(\mathbf{w}_\epsilon(\dot{\mathbf{x}})) = \mathrm{diag}(w_1, w_2, \cdots, w_n), \\ &w_i = 1/\max\{|\dot{x}_i|, \epsilon\}, \quad i = 1, 2, \cdots, n \end{aligned} \quad (9)$$

ensuring that the relationship $\nabla J_\epsilon(\dot{x}_i) = w_i\dot{x}_i$ holds. The procedures are summarized in the Algorithm 1.

After presenting the model update process, we further conduct its convergence analysis. In the real domain, the convergence conditions are established under the NSP assumption [5]. Similarly, this analysis can be extended to the quaternion domain. Before discussing the property, we first introduce the notations used. $[n]$ denotes the set of integers from 1 to n. s is sparsity and S is support set, which refers to the index set corresponding to the non-zero elements. Λ is a subset of S and $\Lambda \cap \Lambda^c = [n]$. $|\Lambda|$ represents the number of elements in the set Λ. Denote $\ker(\dot{\mathbf{A}})$ as $\ker(\dot{\mathbf{A}}) = \left\{\dot{\mathbf{v}} \mid \dot{\mathbf{A}}\dot{\mathbf{v}} = \dot{0}\right\}$.

Algorithm 1 QBP via IRLS

Input: $\dot{\mathbf{A}} \in \mathbb{H}^{m\times n}$, $\dot{\mathbf{y}} \in \mathbb{H}^m$.
Initialization: $\mathbf{w}_0 = \mathbf{1} \in \mathbb{R}^n$, $\epsilon = \infty$, maxIter= 50.

```
while k < maxIter do
    Update x^(k+1) by Eq. (7).
    Update ε^(k+1) by Eq. (8).
    Update W^(k+1) by Eq. (9).
    k = k + 1.
end while
```

Output: $\hat{\dot{\mathbf{x}}} \in \mathbb{H}^n$.

Definition 1. ***Quaternion Null Space Property (QNSP).*** *A matrix $\dot{\mathbf{A}} \in \mathbb{H}^{m\times n}$ is said to satisfy the QNSP of order $s \in \mathbb{N}$ with constant $0 < \rho_s < 1$ if for any set Λ of cardinality $|\Lambda| \le s$, it holds that $\|\dot{\mathbf{v}}_\Lambda\|_1 \le \rho_s\|\dot{\mathbf{v}}_{\Lambda^c}\|_1$, for all $\dot{\mathbf{v}} \in \ker(\dot{\mathbf{A}})$.*

Lemma 1. *Let $\dot{\mathbf{x}}^* = [\dot{x}_1^*, \dot{x}_2^*, \cdots, \dot{x}_n^*]^T \in \mathbb{H}^n$ be an s-sparse vector and set $\dot{\mathbf{y}} = \dot{\mathbf{A}}\dot{\mathbf{x}}^*$. If $0 < \rho_{\hat{s}} < 1 - \frac{2}{\hat{s}+2}$ and $\hat{s} > s + \frac{2\rho_{\hat{s}}}{1-\rho_{\hat{s}}}$, then there exist a constant $k^* \in \mathbb{N}_+$, $\zeta \in \mathbb{R}_+$ such that*

$$\zeta := \frac{\|\dot{\mathbf{x}}^{(k^*)} - \dot{\mathbf{x}}\|_1}{\min\left\{|\dot{x}_i^*|\,\middle|\, i \in S\right\}} < 1, \quad (10)$$

then $\forall k > k^*$, the iterations of QBP satisfies

$$\|\dot{\mathbf{x}}^{(k+1)} - \dot{\mathbf{x}}^*\|_1 \le \frac{\rho_{\hat{s}}(1+\rho_{\hat{s}})}{1-\zeta}\left(1 + \frac{1}{\hat{s} - 1 - s}\right)\|\dot{\mathbf{x}}^{(k)} - \dot{\mathbf{x}}^*\|_1. \quad (11)$$

Remark 1. *The convergence depends on the condition $\zeta := \frac{\|\dot{\mathbf{x}}^{(k^*)} - \dot{\mathbf{x}}^*\|_1}{\min\left\{|\dot{x}_i^*|\,\middle|\, i\in S\right\}} < 1$. The error at each iteration satisfies $\|\dot{\mathbf{x}}^{(k+1)} - \dot{\mathbf{x}}^*\|_1 \le \mu\|\dot{\mathbf{x}}^{(k)} - \dot{\mathbf{x}}^*\|_1$, where $\mu < 1$.*

This property demonstrates the convergence between the previous and the subsequent iterations. Similarly, this result has already been proven in the real domain [8], and the proof can be extended to the quaternion domain. Due to space limitations, readers can refer to the proof in the real domain. The result in the real domain is a special case of ours when the coefficients of the imaginary parts are zero.

4 Experiments

In this section, we first evaluate the performance of the QBP algorithm in signal recovery experiments. And then present a visualization of the convergence process to support Lemma 1.

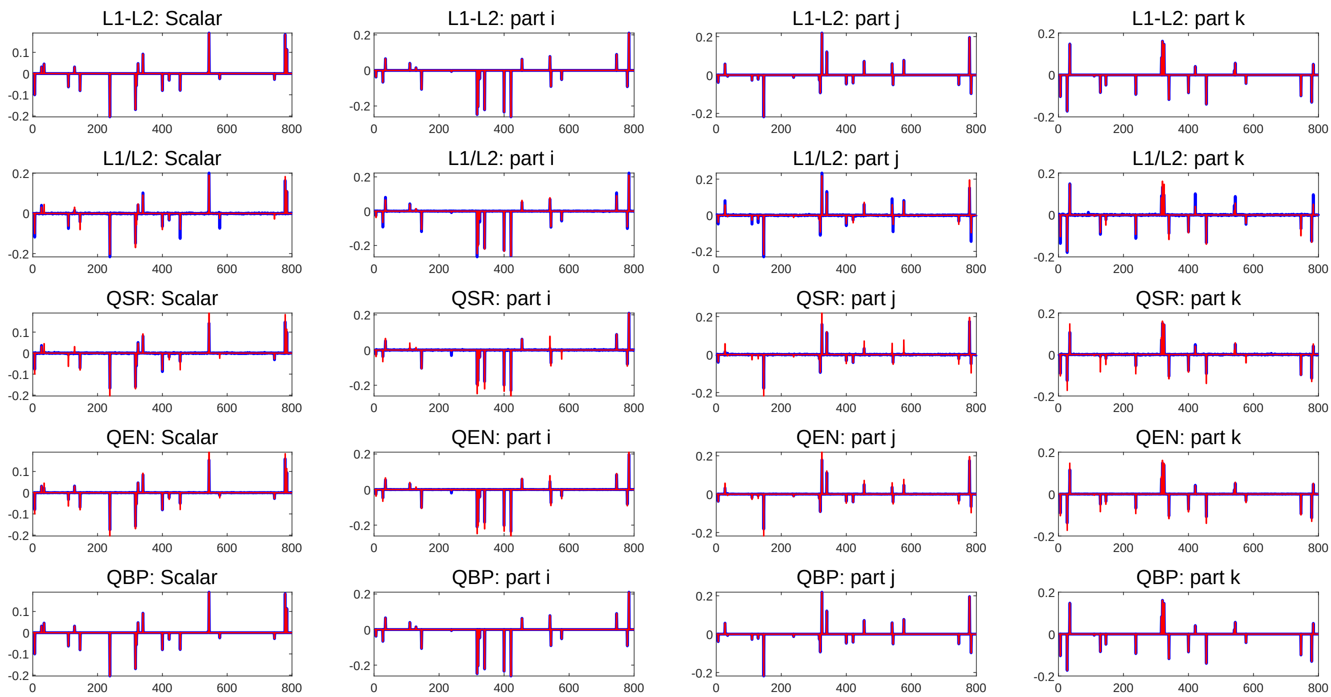

FIGURE 1. Recovery results of signal recovery. The blue lines are the ground-truth, and the red lines represent recovered results of L1-L2 (RelErr = 1.0050e-04), L1/L2 (RelErr = 0.2161), QSR (RelErr = 0.2597), QEN (RelErr = 0.1604) and QBP (RelErr = **4.8460e-08**). From left to right are the scalar and three imagery parts for each quaternion signal.

4.1 Signal recovery

We now evaluate the performance of quaternion basis pursuit by applying it to signal recovery.

Setting: Signal recovery problem is to recover the signal $\hat{\dot{\mathbf{x}}}$ from $\dot{\mathbf{y}} = \dot{\mathbf{A}}\dot{\mathbf{x}}$. $\dot{\mathbf{A}} \in \mathbb{H}^{m\times n}$ is measurement matrix followed by $\mathcal{N}(0, 1/m)$ with $m = \lfloor 2s\log(n/s)\rfloor$, which is known to satisfy the QNSP of order s with high probability [20]. Randomly generate a ground-truth $\dot{\mathbf{x}}^* \in \mathbb{H}^{800}$ with sparsity $s = 20$. $\dot{\mathbf{x}}^*$ satisfies $\sum_{i_j} \dot{x}_{i_j}^2 = 1$ with $i_j \in S \subset [n], j = 1, 2, \cdots, s$.

Comparison methods: The competing methods are quaternion related QSR ($\lambda = 0.1$) [21], QEN ($\lambda_1 = 0.1, \lambda_2 = 0.01$) [13] as well as the real-valued methods L1-L2 [10] and L1/L2 [11]. All the above methods are variations of the ℓ_1 norm. For evaluation, the relative error (RelErr) is employ ed as $\|\hat{\dot{\mathbf{x}}} - \dot{\mathbf{x}}^*\|_2/\|\dot{\mathbf{x}}^*\|_2$.

Results: When $s = 20$, as shown in the Fig. 1, the blue lines represent the ground-truth signals, while the red lines represent the recovered signals obtained by different methods, with the relative error (RelErr) used to quantify the recovery accuracy. QBP achieves the lowest recovery error across all components, almost perfectly matching the ground-truth signals, demonstrating its superiority in quaternion signal recovery. Overall, QBP significantly outperforms the other approaches in terms of recovery accuracy, enabling more precise quaternion signal reconstruction while effectively reducing recovery errors, thus

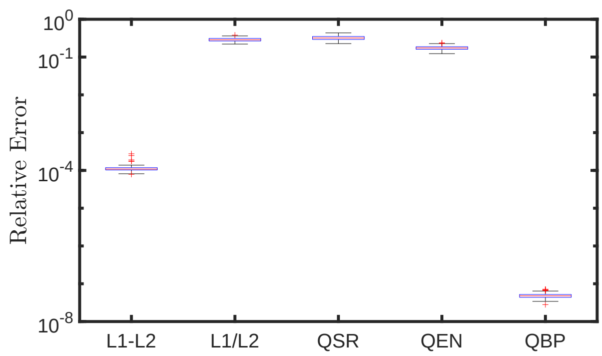

FIGURE 2. The relative errors and variances of signal recovery for different methods over 100 random runs.

showcasing greater robustness and recovery capability.

To further eliminate the impact of randomness on the results, Fig. 2 presents the related error results obtained from 100 experimental runs for each method. QBP outperforms real-valued methods because it can fully exploit the correlation between different channels. QBP outperforms quaternion-based methods because it employs the IRLS update strategy, which enhances the accuracy of estimation.

4.2 Convergence phase

This experiment aims to investigate the convergence phase of the optimization algorithm. Using a simple example, it analyzes how the algorithm exhibits convergence behavior and evaluates its stability across different experimental settings.

Setting: We generate $\dot{\mathbf{x}}^* \in \mathbb{H}^{800}$ with sparsity $s = 20$, other settings are the same as above.

Results: In Fig. 3, the decay of the ℓ_1 error $\|\dot{\mathbf{x}}^{(k)} - \dot{\mathbf{x}}^*\|_1$ for the iterates $\dot{\mathbf{x}}^{(k)}$ produced by Algorithm 1 is analyzed by observing the values of $\zeta^{(k)} := \frac{\|\dot{\mathbf{x}}^{(k)}-\dot{\mathbf{x}}^*\|_1}{\min\left\{|\dot{x}_i^*| \,\middle|\, i \in S\right\}}$, represented in red, and the factor $\mu_1^{(k)} = \frac{\|\dot{\mathbf{x}}^{(k)}-\dot{\mathbf{x}}^*\|_1}{\|\dot{\mathbf{x}}^{(k-1)}-\dot{\mathbf{x}}^*\|_1}$, depicted in blue. Observe that the convergence condition (Eq. (10)) is satisfied after $k = 19$ iterations, as shown by the red dashed line in the Fig. 3.

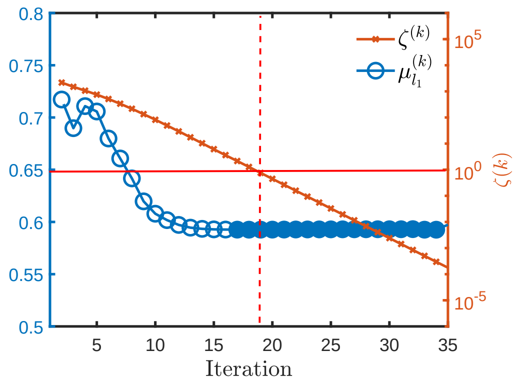

FIGURE 3. Instantaneous linear convergence rates of IRLS for ℓ_1 minimization: Linear convergence factors $\mu_1^{(k)}$ (in blue), filled circles if $S^{(k)} = S$. Error parameter $\zeta^{(k)}$ (in red). Horizontal (red) line: Threshold $\zeta = 1$. Vertical (red) line: $k^* = 19$, the smallest k such that $\zeta^{(k)} < 1$.

We generate $\dot{\mathbf{x}}^* \in \mathbb{H}^{1600}$ with sparsity $s = 40$, other settings remain the same. In the Fig. 4, add the linear convergence factor $\mu^{(k)} = \frac{\mathcal{J}_{\epsilon(k)}(\dot{\mathbf{x}}^{(k)})-\|\dot{\mathbf{x}}^*\|_1}{\mathcal{J}_{\epsilon(k-1)}(\dot{\mathbf{x}}^{(k-1)})-\|\dot{\mathbf{x}}^*\|_1}$ that tracks the linear convergence behavior of the smoothed ℓ_1 norm objective $\mathcal{J}$ defined in (4). In addition to the observations in Fig. 4, it is noticed that $\mu^{(k)}$

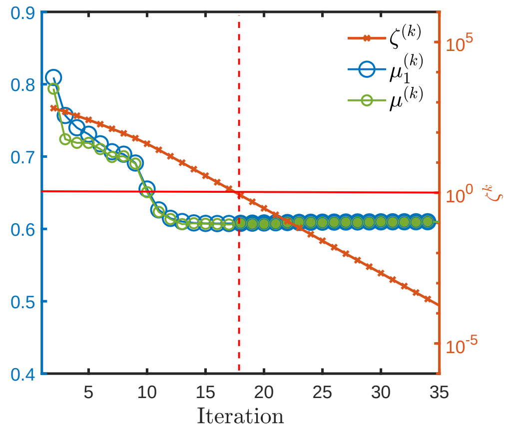

FIGURE 4. Instantaneous linear convergence rates for ℓ_1 minimization: Linear convergence factors $\mu_1^{(k)}$ (in blue) and $\mu^{(k)}$ (in green), filled circles if $S^{(k)} = S$. Error parameter $\zeta^{(k)}$ (in red). Horizontal (red) line: Threshold $\zeta = 1$. Vertical (red) line: $k^* = 18$, the smallest k such that $\zeta^{(k)} < 1$.

and $\mu_1^{(k)}$ exhibit a very similar behavior. This further validates the effectiveness of the smoothing approach in preserving the expected convergence properties.

5 Conclusion

This paper extends the Basis Pursuit algorithm to the quaternion domain and proposes a Quaternion Basis Pursuit (QBP) algorithm to address the sparse optimization needs of high-dimensional data. To solve the sparse optimization problem in the quaternion domain, we adopt the Iteratively Reweighted Least Squares (IRLS) method. This ensures that the algorithm can start from any initial point and achieve linear convergence to the optimal solution, thereby avoiding the local minima issues encountered by traditional methods. Furthermore, we designed and performed a series of numerical experiments, and the results confirm the effectiveness of the proposed algorithm.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant Nos. 62276111, 62076041 and 61806027. This work was supported by Research Grants of Huazhong Agricultural University of Grant No. 2662024XXPY005.

References

[1] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?," *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[2] G. Davis, S. Mallat, and M. Avellaneda, "Adaptive greedy approximations," *Constructive approximation*, vol. 13, pp. 57–98, 1997.

[3] D. L. Donoho and B. F. Logan, "Signal recovery and the large sieve," *SIAM Journal on Applied Mathematics*, vol. 52, no. 2, pp. 577–591, 1992.

[4] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM review*, vol. 43, no. 1, pp. 129–159, 2001.

[5] E. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.

[6] R. Chartrand and W. Yin, "Iteratively reweighted algorithms for compressive sensing," in *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3869–3872, 2008.

[7] L. Peng, C. Kümmerle, and R. Vidal, "Global linear and local superlinear convergence of irls for non-smooth robust regression," *Advances in neural information processing systems*, vol. 35, pp. 28972–28987, 2022.

[8] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Communications on Pure and Applied Mathematics: A Journal Issued by the Courant Institute of Mathematical Sciences*, vol. 63, no. 1, pp. 1–38, 2010.

[9] C. Wang, Z. Li, and R. P. Agarwal, "A new quaternion hyper-complex space with hyper argument and basic functions via quaternion dynamic equations," *The Journal of Geometric Analysis*, vol. 32, no. 2, p. 67, 2022.

[10] Y. Lou, P. Yin, Q. He, and J. Xin, "Computing sparse representation in a highly coherent dictionary based on difference of l_1 and l_2," *Journal of Scientific Computing*, vol. 64, pp. 178–196, 2015.

[11] Y. Rahimi, C. Wang, H. Dong, and Y. Lou, "A scale-invariant approach for sparse signal recovery," *SIAM Journal on Scientific Computing*, vol. 41, no. 6, pp. A3649–A3672, 2019.

[12] 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.

[13] S.-Y. Wang and C.-M. Zou, "Quaternion elastic net for signal recovery," in *2023 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)*, pp. 32–37, IEEE, 2023.

[14] J. F. Mota, J. M. Xavier, P. M. Aguiar, and M. Puschel, "Distributed basis pursuit," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1942–1956, 2011.

[15] E. Van Den Berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *Siam journal on scientific computing*, vol. 31, no. 2, pp. 890–912, 2009.

[16] J.-H. Chang, J.-J. Ding, *et al.*, "Quaternion matrix singular value decomposition and its applications for color image processing," in *Proceedings 2003 international conference on image processing (Cat. No. 03CH37429)*, vol. 1, pp. I–805, IEEE, 2003.

[17] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Transactions on Signal Processing*, vol. 65, no. 3, pp. 794–816, 2016.

[18] C. Kümmerle, C. Mayrink Verdun, and D. Stöger, "Iteratively reweighted least squares for basis pursuit with global linear convergence rate," *Advances in neural information processing systems*, vol. 34, pp. 2873–2886, 2021.

[19] P. W. Holland and R. E. Welsch, "Robust regression using iteratively reweighted least-squares," *Communications in Statistics-theory and Methods*, vol. 6, no. 9, pp. 813–827, 1977.

[20] J. A. Tropp, "A mathematical introduction to compressive sensing book review," *Bulletin of the American Mathematical Society*, vol. 54, no. 1, pp. 151–165, 2017.

[21] 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.