Quaternion Reweighted Low-rank Factorization with Deep Denoising Prior for Color Image Inpainting (Supplementary Material)

Zhijie Wang, Liangtian He*, Shaobing Gao, Jifei Miao, Liang-Jian Deng, Jun Liu

Appendix A. Quaternion algebra

As an extension of the complex space $\mathbb{C}$, the quaternion space $\mathbb{H}$ was first introduced by Hamilton [1]. A quaternion number $\dot{a} \in \mathbb{H}$, which consists of a real part and three imaginary parts, is typically expressed in the form:

$$\dot{a} = a_0 + a_1\mathtt{i} + a_2\mathtt{j} + a_3\mathtt{k}, \tag{A.1}$$

where $a_0, a_1, a_2, a_3 \in \mathbb{R}$, and $\mathtt{i}, \mathtt{j}, \mathtt{k}$ are the three imaginary units obeying the following rules:

$$\begin{cases} \mathtt{i}^2 = \mathtt{j}^2 = \mathtt{k}^2 = \mathtt{ijk} = -1, \\ \mathtt{ij} = \mathtt{k} = -\mathtt{ji}, \mathtt{jk} = \mathtt{i} = -\mathtt{kj}, \mathtt{ki} = \mathtt{j} = -\mathtt{ik}. \end{cases} \tag{A.2}$$

In particular, if the real part $a_0 := \Re(\dot{a}) = 0$, then $\dot{a} = a_1\mathtt{i} + a_2\mathtt{j} + a_3\mathtt{k}$ is a pure quaternion. It is important to note that, in general, the multiplication of two quaternions does not follow the commutative property, i.e., $\dot{a}\dot{b} \neq \dot{b}\dot{a}$. The conjugate and modulus of $\dot{a}$ are represented as $\dot{a}^* = a_0 - a_1\mathtt{i} - a_2\mathtt{j} - a_3\mathtt{k}$ and $|\dot{a}| = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2}$, respectively.

Analogously, given a quaternion matrix $\dot{\mathbf{A}} = (\dot{a}_{ij}) \in \mathbb{H}^{M\times N}$ that be written as $\dot{\mathbf{A}} = \mathbf{A}_0 + \mathbf{A}_1\mathtt{i} + \mathbf{A}_2\mathtt{j} + \mathbf{A}_3\mathtt{k}$, where $\mathbf{A}_s \in \mathbb{R}^{M\times N}(s = 0, 1, 2, 3)$. If $\mathbf{A}_0 = 0$, $\dot{\mathbf{A}}$ is called a pure quaternion matrix. The conjugate, transpose and conjugate transpose of $\dot{\mathbf{A}}$ are expressed as $\dot{\mathbf{A}}^* = (\dot{a}^*_{mn})$, $\dot{\mathbf{A}}^T = (\dot{a}_{nm})$, $\dot{\mathbf{A}}^H = (\dot{a}^*_{nm})$, respectively, where $1 \le m \le M$, $1 \le n \le N$. The Frobenius norm of $\dot{\mathbf{A}}$ is defined as $||\dot{\mathbf{A}}||_F = \sqrt{\sum_{m=1}^{M}\sum_{n=1}^{N} |\dot{a}_{mn}|^2} = \sqrt{tr(\dot{\mathbf{A}}^H\dot{\mathbf{A}})}$. In particular, according to the Cayley-Dickson notation [2], $\dot{\mathbf{A}}$ can be expressed as $\dot{\mathbf{A}} = \mathbf{A}_a + \mathbf{A}_b\mathtt{j}$, where

*Corresponding author.
Email addresses: wzjAUST@163.com (Zhijie Wang), helt@ahu.edu.cn (Liangtian He), gaoshaobing@scu.edu.cn (Shaobing Gao), jifmiao@163.com (Jifei Miao), liangjian.deng@uestc.edu.cn (Liang-Jian Deng), liuj292@nenu.edu.cn (Jun Liu)

$\mathbf{A}_a = \mathbf{A}_0 + \mathbf{A}_1\mathtt{i} \in \mathbb{C}^{M\times N}$, $\mathbf{A}_b = \mathbf{A}_2 + \mathbf{A}_3\mathtt{i} \in \mathbb{C}^{M\times N}$. Then, the quaternion matrix $\dot{\mathbf{A}}$ can be transformed into the complex adjoint matrix:

$$\mathbf{A}_c = \begin{pmatrix} \mathbf{A}_a & \mathbf{A}_b \\ -\mathbf{A}_b^* & \mathbf{A}_a^* \end{pmatrix} \in \mathbb{C}^{2M\times 2N}. \tag{A.3}$$

Based on (A.3), the QSVD of the quaternion matrix $\dot{\mathbf{A}}$ can be computed by the classical complex SVD on the complex matrix $\mathbf{A}_c$. Specifically, the QSVD of $\dot{\mathbf{A}} \in \mathbb{H}^{M\times N}$ ($\dot{\mathbf{A}} = \dot{\mathbf{U}}\mathbf{D}\dot{\mathbf{V}}^H$) and the SVD of $\mathbf{A}_c \in \mathbb{C}^{2M\times 2N}$ ($\mathbf{A}_c = \mathbf{U}\check{\mathbf{D}}\mathbf{V}^H$) has the following relation:

$$\begin{cases} \mathbf{D} = \mathrm{row}_{odd}(\mathrm{col}_{odd}(\check{\mathbf{D}})), \\ \dot{\mathbf{U}} = \mathrm{col}_{odd}(\mathbf{U}_1) + \mathrm{col}_{odd}(-(\mathbf{U}_2)^*)\mathtt{j}, \\ \dot{\mathbf{V}} = \mathrm{col}_{odd}(\mathbf{V}_1) + \mathrm{col}_{odd}(-(\mathbf{V}_2)^*)\mathtt{j}, \end{cases} \tag{A.4}$$

where

$$\mathbf{U} = \begin{pmatrix} (\mathbf{U}_1)_{M\times 2M} \\ (\mathbf{U}_2)_{M\times 2M} \end{pmatrix}, \mathbf{V} = \begin{pmatrix} (\mathbf{V}_1)_{N\times 2N} \\ (\mathbf{V}_2)_{N\times 2N} \end{pmatrix}, \tag{A.5}$$

and $\mathrm{row}_{odd}(\mathbf{X})$ and $\mathrm{col}_{odd}(\mathbf{X})$ denote the extraction of the odd-numbered rows and columns of $\mathbf{X}$, respectively.

In the following, we present several essential definitions and propositions related to quaternion matrices.

Definition 1. *(Quaternion Rank [3]) Given a quaternion matrix $\dot{\mathbf{A}} \in \mathbb{H}^{M\times N}$, its rank, denoted by r, is defined as the number of its nonzero quaternion singular values of $\dot{\mathbf{A}}$.*

Definition 2. *(Quaternion Nuclear Norm (QNN) [3]) Given a quaternion matrix $\dot{\mathbf{A}} \in \mathbb{H}^{M\times N}$ with $\mathrm{rank}(\dot{\mathbf{A}}) = r$, the QNN of $\dot{\mathbf{A}}$ denoted as $||\dot{\mathbf{A}}||_*$, is defined as the sum of its singular values $\sigma_i(\dot{\mathbf{A}})$, i.e., $||\dot{\mathbf{A}}||_* = \sum_{i=1}^r \sigma_i(\dot{\mathbf{A}})$.*

Proposition 1. *(Quaternion Singular Value Decomposition (QSVD) [4]) Given a quaternion matrix $\dot{\mathbf{A}} \in \mathbb{H}^{M\times N}$ with $\mathrm{rank}(\dot{\mathbf{A}}) = r$, there exists two unitary quaternion matrices $\dot{\mathbf{U}} \in \mathbb{H}^{M\times M}$ and $\dot{\mathbf{V}} \in \mathbb{H}^{N\times N}$ such that*

$$\dot{\mathbf{A}} = \dot{\mathbf{U}} \begin{pmatrix} \mathbf{\Sigma}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \dot{\mathbf{V}}^H, \tag{A.6}$$

where $\mathbf{\Sigma}_r = diag\{\sigma_1, \sigma_2, \ldots, \sigma_r\} \in \mathbb{R}^{r\times r}$ and σ_i $(i = 1, \ldots, r)$ are the singular values of $\dot{\mathbf{A}}$.

Proposition 2. *(Quaternion Singular Value Thresholding (QSVT) [3]) Given a quaternion matrix $\dot{\mathbf{A}} \in \mathbb{H}^{M\times N}$ and a real number $\lambda > 0$, for the following optimization problem:*

$$\arg\min_{\dot{\mathbf{A}}} \lambda||\dot{\mathbf{A}}||_* + \frac{1}{2}||\dot{\mathbf{A}} - \dot{\mathbf{B}}||_F^2. \tag{A.7}$$

The optimum can be obtained by the QSVT operation:

$$\dot{\mathbf{A}}^{\star} = \dot{\mathbf{U}} S_{\lambda}(\mathbf{\Sigma}) \dot{\mathbf{V}}^{H}, \tag{A.8}$$

where $\dot{\mathbf{U}}$, $\mathbf{\Sigma}$, $\dot{\mathbf{V}}$ *are the components obtained from the QSVD of* $\dot{\mathbf{B}}$, *and the operator* $S_{\lambda}(\mathbf{\Sigma}) = diag\{\max(\sigma_i(\dot{\mathbf{B}}) - \lambda, 0)\}$.

Proposition 3. *(Quaternion Bilinear Factorization (QBF) [5]) Given a quaternion matrix* $\dot{\mathbf{A}} \in \mathbb{H}^{M \times N}$, *if* $\operatorname{rank}(\dot{\mathbf{A}}) = r \leq d$, *then* $\dot{\mathbf{A}}$ *can be equivalently rewritten into the product of two smaller quaternion matrices* $\dot{\mathbf{A}} = \dot{\mathbf{U}}\dot{\mathbf{V}}^{H}$, *where* $\dot{\mathbf{U}} \in \mathbb{H}^{M \times d}$ *and* $\dot{\mathbf{V}} \in \mathbb{H}^{N \times d}$ *satisfy* $\operatorname{rank}(\dot{\mathbf{U}}) = \operatorname{rank}(\dot{\mathbf{V}}) = r$.

Proposition 4. *(Von Neumanns Trace Inequality [5]) Given two quaternion matrices* $\dot{\mathbf{A}}, \dot{\mathbf{B}} \in \mathbb{H}^{M \times N}$, *the following equality holds:*

$$\Re(tr(\dot{\mathbf{A}}^{H}\dot{\mathbf{B}})) \leq \Re(tr\left(\sigma(\dot{\mathbf{A}})^{H}\sigma(\dot{\mathbf{B}})\right), \tag{A.9}$$

where $\sigma(\dot{\mathbf{A}})$ *and* $\sigma(\dot{\mathbf{B}})$ *denote the ordered singular values of* $\dot{\mathbf{A}}$ *and* $\dot{\mathbf{B}}$ *respectively, under the same ordering.*

Appendix B. The proof of Theorem 1

We first introduce the following lemma, which serves as the cornerstone of the subsequent proof.

Lemma 1. *Let* $\dot{\mathbf{U}} \in \mathbb{H}^{M \times r}$, $\dot{\mathbf{V}} \in \mathbb{H}^{N \times r}$ *be two quaternion matrices, and let* $g(\cdot) : \mathbb{R} \to \mathbb{R}$ *be a function such that* $\varphi(x) = g(e^{x})$ *is convex and increasing on the interval* $[\min\{\sigma_l(\dot{\mathbf{U}}\dot{\mathbf{V}}^{H}), \sigma_l(\dot{\mathbf{U}})\sigma_l(\dot{\mathbf{V}})\}, \sigma_1(\dot{\mathbf{U}})\sigma_1(\dot{\mathbf{V}})]$, *where* $l = \min\{M, N, r\}$. *Then, for all* $j = 1, 2, \cdots, l$, *the following inequality holds:*

$$\sum_{i=1}^{j} g\left(\sigma_i(\dot{\mathbf{U}}\dot{\mathbf{V}}^{H})\right) \leq \sum_{i=1}^{j} g\left(\sigma_i(\dot{\mathbf{U}})\sigma_i(\dot{\mathbf{V}})\right).$$

Proof. According to the transformed complex adjoint matrix of quaternion matrix, we have:

$$\begin{aligned}\sum_{i=1}^{j} g\left(\sigma_i(\dot{\mathbf{U}}\dot{\mathbf{V}}^{H})\right) &= \frac{1}{2}\sum_{i=1}^{j} g\left(\sigma_i(\mathbf{U}\mathbf{V}^{H})_c\right) = \frac{1}{2}\sum_{i=1}^{j} g\left(\sigma_i(\mathbf{U}_c(\mathbf{V}^{H})_c)\right) \\ &\leq \frac{1}{2}\sum_{i=1}^{j} g\left(\sigma_i(\mathbf{U}_c)\sigma_i((\mathbf{V}^{H})_c)\right) = \sum_{i=1}^{j} g\left(\sigma_i(\dot{\mathbf{U}})\sigma_i(\dot{\mathbf{V}})\right),\end{aligned}$$

where the first and second equalities can be directly verified based on (A.3), (A.4) and (A.5) (or refer to [4], and the inequality follows from [6] (Theorem 3.3.14). ∎

According to Lemma 1, by defining $g(x) = c(x+\epsilon)^{p-1}x$ with $0 < p \leq 1$, it is evident that $\varphi(x) = c(e^x+\epsilon)^{p-1}e^x$ is convex and increasing on the corresponding interval. Applying a special case of $j = \min\{M, N, r\}$ to Lemma 1 leads to the following corollary.

Corollary 1. *For two quaternion matrices $\dot{\mathbf{U}} \in \mathbb{H}^{M\times r}, \dot{\mathbf{V}} \in \mathbb{H}^{N\times r}$ and $0 < p \leq 1$, the following inequality holds:*

$$\sum_{i=1}^{\min\{M,N,r\}} \alpha_i \sigma_i(\dot{\mathbf{U}}\dot{\mathbf{V}}^H) \leq \sum_{i=1}^{\min\{M,N,r\}} \hat{\alpha}_i \left(\sigma_i(\dot{\mathbf{U}})\sigma_i(\dot{\mathbf{V}})\right),$$

where the scalars

$$\alpha_i = c\left(\sigma_i(\dot{\mathbf{U}}\dot{\mathbf{V}}^H) + \epsilon\right)^{p-1}, \quad \hat{\alpha}_i = c\left(\sigma_i(\dot{\mathbf{U}})\sigma_i(\dot{\mathbf{V}}) + \epsilon\right)^{p-1}, \quad i = 1, \cdots, r.$$

Next, we incorporate Corollary 1 to derive the upper bound of the proposed quaternion reweighted nuclear norm. Without loss of generality, we consider two distinct cases: $r \leq \min\{M, N\}$ and $r > \min\{M, N\}$.

1) In the first case, where $r \leq \min\{M, N\}$, the following inequalities hold:

$$\begin{aligned}||\dot{\mathbf{X}}||_{\mathbf{w},*} &= \sum_{i=1}^{r} \alpha_i \sigma_i(\dot{\mathbf{X}}) = \sum_{i=1}^{r} \alpha_i \sigma_i(\dot{\mathbf{U}}\dot{\mathbf{V}}^H) \leq \sum_{i=1}^{r} \hat{\alpha}_i \left(\sigma_i(\dot{\mathbf{U}})\sigma_i(\dot{\mathbf{V}})\right) \\ &\leq \frac{1}{2}\sum_{i=1}^{r} \hat{\alpha}_i \left(\sigma_i^2(\dot{\mathbf{U}}) + \sigma_i^2(\dot{\mathbf{V}})\right) = \frac{1}{2}\left(||\dot{\mathbf{U}}||_{\mathbf{w},F}^2 + ||\dot{\mathbf{V}}||_{\mathbf{w},F}^2\right).\end{aligned}$$

Suppose the QSVD of $\dot{\mathbf{X}}$ is given by $\dot{\mathbf{X}} = \dot{\mathbf{L}}\mathbf{\Sigma}\dot{\mathbf{R}}^H$. We define $\dot{\mathbf{U}}^\star = \dot{\mathbf{L}}\mathbf{\Sigma}^{\frac{1}{2}}\dot{\mathbf{E}}^H$ and $\dot{\mathbf{V}}^\star = \dot{\mathbf{E}}\mathbf{\Sigma}^{\frac{1}{2}}\dot{\mathbf{R}}^H$, where $\dot{\mathbf{E}}$ is a unitary quaternion matrix. Then, it holds that $\dot{\mathbf{X}} = \dot{\mathbf{U}}^\star(\dot{\mathbf{V}}^\star)^H$, and we obtain the equality:

$$||\dot{\mathbf{X}}||_{\mathbf{w},*} = \frac{1}{2}\left(||\dot{\mathbf{U}}^\star||_{\mathbf{w},F}^2 + ||\dot{\mathbf{V}}^\star||_{\mathbf{w},F}^2\right).$$

2) In the second case, where $r > \min\{M, N\}$, we have:

$$\begin{aligned}||\dot{\mathbf{X}}||_{\mathbf{w},*} &= \sum_{i=1}^{\min\{M,N\}} \alpha_i \sigma_i(\dot{\mathbf{X}}) \leq \sum_{i=1}^{\min\{M,N\}} \hat{\alpha}_i \sigma_i(\dot{\mathbf{U}})\sigma_i(\dot{\mathbf{V}}) \\ &\leq \frac{1}{2}\sum_{i=1}^{\min\{M,N\}} \hat{\alpha}_i \left(\sigma_i^2(\dot{\mathbf{U}}) + \sigma_i^2(\dot{\mathbf{V}})\right) = \frac{1}{2}\left(||\dot{\mathbf{U}}||_{\mathbf{w},F}^2 + ||\dot{\mathbf{V}}||_{\mathbf{w},F}^2\right).\end{aligned}$$

Again, we use the QSVD of $\dot{\mathbf{X}} = \dot{\mathbf{L}}\mathbf{\Sigma}\dot{\mathbf{R}}^H$ to define $\dot{\mathbf{U}}^\star = \dot{\mathbf{L}}_{\mathrm{a}}\mathbf{\Sigma}_{\mathrm{a}}^{\frac{1}{2}}\dot{\mathbf{E}}^H$ and $\dot{\mathbf{V}}^\star = \dot{\mathbf{E}}\mathbf{\Sigma}_{\mathrm{a}}^{\frac{1}{2}}\dot{\mathbf{R}}_{\mathrm{a}}^H$, where $\dot{\mathbf{L}}_{\mathrm{a}}$ and $\dot{\mathbf{R}}_{\mathrm{a}}$ are constructed from $\dot{\mathbf{L}}$ and $\dot{\mathbf{R}}$ by appending $(r - \min\{M, N\})$ zero columns to $\dot{\mathbf{L}}$ and $\dot{\mathbf{R}}$, respectively, and $\mathbf{\Sigma}_{\mathrm{a}}$ is also constructed similarly from $\mathbf{\Sigma}$ through the zero padding. In this way, $\dot{\mathbf{X}} = \dot{\mathbf{U}}^\star(\dot{\mathbf{V}}^\star)^H$ holds, thereby completing the proof.

Appendix C. The proof of Theorem 2

According to Theorem 1, we have:

$$\begin{aligned}&\min_{\dot{\mathbf{U}},\dot{\mathbf{V}}} F_1(\dot{\mathbf{U}},\dot{\mathbf{V}}) \triangleq \frac{\lambda}{2}(||\dot{\mathbf{U}}||_{\mathbf{w},F}^2 + ||\dot{\mathbf{V}}||_{\mathbf{w},F}^2) + \frac{1}{2}||\mathcal{P}_\Omega(\dot{\mathbf{U}}\dot{\mathbf{V}}^H - \dot{\mathbf{Y}})||_F^2 \\ =&\min_{\dot{\mathbf{U}},\dot{\mathbf{V}}} F_2(\dot{\mathbf{U}},\dot{\mathbf{V}}) \triangleq \lambda||\dot{\mathbf{U}}\dot{\mathbf{V}}^H||_{\mathbf{w},*} + \frac{1}{2}||\mathcal{P}_\Omega(\dot{\mathbf{U}}\dot{\mathbf{V}}^H - \dot{\mathbf{Y}})||_F^2,\end{aligned} \tag{C.1}$$

which can be proved by contradiction as follows.

Assume that $\min_{\dot{\mathbf{U}}_1,\dot{\mathbf{V}}_1} F_1(\dot{\mathbf{U}}_1,\dot{\mathbf{V}}_1) < \min_{\dot{\mathbf{U}}_2,\dot{\mathbf{V}}_2} F_2(\dot{\mathbf{U}}_2,\dot{\mathbf{V}}_2)$. Then, we can construct a compared solution $F_2(\dot{\mathbf{U}}_1,\dot{\mathbf{V}}_1)$, which leads to $\min_{\dot{\mathbf{U}}_1,\dot{\mathbf{V}}_1} F_1(\dot{\mathbf{U}}_1,\dot{\mathbf{V}}_1) \geqslant F_2(\dot{\mathbf{U}}_1,\dot{\mathbf{V}}_1)$ based on Theorem 1, thus contradicting with the assumption.

Conversely, assume that $\min_{\dot{\mathbf{U}}_1,\dot{\mathbf{V}}_1} F_1(\dot{\mathbf{U}}_1,\dot{\mathbf{V}}_1) > \min_{\dot{\mathbf{U}}_2,\dot{\mathbf{V}}_2} F_2(\dot{\mathbf{U}}_2,\dot{\mathbf{V}}_2)$, we can also construct a compared solution $F_1(\dot{\mathbf{U}}^\star,\dot{\mathbf{V}}^\star)$, where $\dot{\mathbf{U}}^\star = \dot{\mathbf{L}}\mathbf{\Sigma}^{\frac{1}{2}}\dot{\mathbf{E}}^H$, $\dot{\mathbf{V}}^\star = \dot{\mathbf{E}}\mathbf{\Sigma}^{\frac{1}{2}}\dot{\mathbf{R}}^H$, $\dot{\mathbf{E}}$ is a unitary quaternion matrix, and $\dot{\mathbf{U}}_2\dot{\mathbf{V}}_2^H = \dot{\mathbf{L}}\mathbf{\Sigma}\dot{\mathbf{R}}^H$ denotes the QSVD of $(\dot{\mathbf{U}}_2\dot{\mathbf{V}}_2^H)$. By Theorem 1, we have $F_1(\dot{\mathbf{U}}^\star,\dot{\mathbf{V}}^\star) = \min_{\dot{\mathbf{U}}_2,\dot{\mathbf{V}}_2} F_2(\dot{\mathbf{U}}_2,\dot{\mathbf{V}}_2)$, which contradicts with the assumption again and thus (C.1) is proved.

Furthermore, for any quaternion matrix $\dot{\mathbf{X}} \in \mathbb{H}^{M\times N}$ with $\text{rank}(\dot{\mathbf{X}}) = r^\star \leq r$, it can be decomposed into $\dot{\mathbf{X}} = \dot{\mathbf{U}}\dot{\mathbf{V}}^H$ with $\dot{\mathbf{U}} \in \mathbb{H}^{M\times r}$ and $\dot{\mathbf{V}} \in \mathbb{H}^{N\times r}$. Therefore, we have:

$$\begin{aligned}\min_{\dot{\mathbf{X}}} \lambda||\dot{\mathbf{X}}||_{\mathbf{w},*} + \frac{1}{2}||\mathcal{P}_\Omega(\dot{\mathbf{X}} - \dot{\mathbf{Y}})||_F^2 &= \min_{\dot{\mathbf{X}}:\text{rank}(\dot{\mathbf{X}})=r^\star} \lambda||\dot{\mathbf{X}}||_{\mathbf{w},*} + \frac{1}{2}||\mathcal{P}_\Omega(\dot{\mathbf{X}} - \dot{\mathbf{Y}})||_F^2 \\ &= \min_{\dot{\mathbf{U}},\dot{\mathbf{V}}} \lambda||\dot{\mathbf{U}}\dot{\mathbf{V}}^H||_{\mathbf{w},*} + \frac{1}{2}||\mathcal{P}_\Omega(\dot{\mathbf{U}}\dot{\mathbf{V}}^H - \dot{\mathbf{Y}})||_F^2,\end{aligned} \tag{C.2}$$

where the first equality holds since the rank of the solution to (5) is $r^\star$. Combining (C.1) and (C.2) completes the proof of Theorem 2.

Appendix D. The proof of Theorem 3

It is evident that Theorem 3 is equivalent to proving the following formula:

$$\arg\min_{\dot{\mathbf{A}}} F(\dot{\mathbf{A}}) \triangleq \gamma\sum_{i=1}^{s} w_i\sigma_i^2(\dot{\mathbf{A}}) + ||\dot{\mathbf{A}} - \dot{\mathbf{B}}||_F^2, \tag{D.1}$$

where $\dot{\mathbf{B}} = \dot{\mathbf{L}}\mathbf{\Sigma}\dot{\mathbf{R}}^H$ is the QSVD of $\dot{\mathbf{B}}$, and $\mathbf{W} = diag\{w_1, w_2, \cdots, w_s\}$.

Specifically, the objective function can be rewritten as:

$$
\begin{aligned}
F(\dot{\mathbf{A}}) &= \gamma \sum_{i=1}^{s} w_i \sigma_i^2(\dot{\mathbf{A}}) + ||\dot{\mathbf{A}} - \dot{\mathbf{B}}||_F^2 \\
&= \gamma \sum_{i=1}^{s} w_i \sigma_i^2(\dot{\mathbf{A}}) + (||\dot{\mathbf{A}}||_F^2 + ||\dot{\mathbf{B}}||_F^2 - 2\langle \dot{\mathbf{A}}, \dot{\mathbf{B}} \rangle) \\
&= \gamma \sum_{i=1}^{s} w_i \sigma_i^2(\dot{\mathbf{A}}) + \Big(||\dot{\mathbf{A}}||_F^2 + ||\dot{\mathbf{B}}||_F^2 - 2tr(\dot{\mathbf{A}}^H \dot{\mathbf{B}})\Big).
\end{aligned} \tag{D.2}
$$

According to the Von Neumann's Trace Inequality of Proposition 4 in Appendix A, we have:

$$
\begin{aligned}
F(\dot{\mathbf{A}}) &\geqslant \gamma \sum_{i=1}^{s} w_i \sigma_i^2(\dot{\mathbf{A}}) + \Big(||\sigma(\dot{\mathbf{A}})||_F^2 + ||\sigma(\dot{\mathbf{B}})||_F^2 - 2tr(\sigma(\dot{\mathbf{A}})\sigma(\dot{\mathbf{B}}))\Big) \\
&= \gamma \sum_{i=1}^{s} w_i \sigma_i^2(\dot{\mathbf{A}}) + ||\sigma(\dot{\mathbf{A}}) - \sigma(\dot{\mathbf{B}})||_F^2,
\end{aligned} \tag{D.3}
$$

where $\sigma(\dot{\mathbf{A}})$ and $\sigma(\dot{\mathbf{B}})$ are the ordered singular values of $\dot{\mathbf{A}}$ and $\dot{\mathbf{B}}$ with the same order, respectively.

Let $\dot{\mathbf{B}} = \dot{\mathbf{L}} diag\{\sigma(\dot{\mathbf{B}})\}\dot{\mathbf{R}}^H$ be the QSVD of $\dot{\mathbf{B}}$. The equation in (D.3) holds if and only if $\dot{\mathbf{L}}$ and $\dot{\mathbf{R}}$ are the unitary quaternion matrices from the QSVD $\dot{\mathbf{A}} = \dot{\mathbf{L}} diag\{\sigma(\dot{\mathbf{A}})\}\dot{\mathbf{R}}^H$. Subsequently, we investigate the optimal solution of the lower bound in (D.3), that is,

$$
\begin{aligned}
\hat{\sigma}(\dot{\mathbf{A}}) &= \arg\min_{\sigma(\dot{\mathbf{A}})} F(\sigma(\dot{\mathbf{A}})) \\
&= \arg\min_{\sigma(\dot{\mathbf{A}})} \gamma \sum_{i=1}^{s} w_i \sigma_i^2(\dot{\mathbf{A}}) + ||\sigma(\dot{\mathbf{A}}) - \sigma(\dot{\mathbf{B}})||_F^2.
\end{aligned} \tag{D.4}
$$

Let $\frac{\partial F(\delta(\dot{\mathbf{A}}))}{\partial(\sigma(\dot{\mathbf{A}}))} = \dot{\mathbf{0}}$, we have:

$$
\hat{\sigma}(\dot{\mathbf{A}}) = (\gamma \mathbf{W} + \mathbf{I}_s)^{-1}\sigma(\dot{\mathbf{B}}),
$$

where $\mathbf{W} = diag\{w_1, w_2, \cdots, w_s\}$ and thus the optimum is given by:

$$
\begin{aligned}
\dot{\mathbf{A}}^\star = \arg\min_{\dot{\mathbf{A}}} F(\dot{\mathbf{A}}) &= \dot{\mathbf{L}} diag\{\hat{\sigma}(\dot{\mathbf{A}})\}\dot{\mathbf{R}}^H \\
&= \dot{\mathbf{L}}\left((\gamma \mathbf{W} + \mathbf{I}_s)^{-1}\mathbf{\Sigma}\right)\dot{\mathbf{R}}^H.
\end{aligned} \tag{D.5}
$$

Appendix E. The proof of Theorem 4

1. We begin by proving that sequences $\{\dot{\mathbf{Q}}_U^k\}$, $\{\dot{\mathbf{Q}}_V^k\}$, $\{\dot{\mathbf{U}}^k\}$, $\{\dot{\mathbf{V}}^k\}$, and $\{\dot{\mathbf{X}}^k\}$ are Cauchy sequences. The Lagrange multipliers of the QRLMF method are updated according to equation (22), from which it follows that:

$$\begin{cases} \lim\limits_{k\to+\infty} ||\dot{\mathbf{U}}^{k+1} - \dot{\mathbf{Q}}_U^{k+1}||_F = 0, \\ \lim\limits_{k\to+\infty} ||\dot{\mathbf{V}}^{k+1} - \dot{\mathbf{Q}}_V^{k+1}||_F = 0, \\ \lim\limits_{k\to+\infty} ||\dot{\mathbf{X}}^{k+1} - \dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H||_F = 0. \end{cases} \tag{E.1}$$

Therefore, the sequence $\{(\dot{\mathbf{U}}^k, \dot{\mathbf{V}}^k, \dot{\mathbf{Q}}_U^k, \dot{\mathbf{Q}}_V^k, \dot{\mathbf{X}}^k)\}$ converges to a feasible solution.

To prove that sequences $\{\dot{\mathbf{Q}}_U^k\}$, $\{\dot{\mathbf{Q}}_V^k\}$, $\{\dot{\mathbf{U}}^k\}$, $\{\dot{\mathbf{V}}^k\}$ and $\{\dot{\mathbf{X}}^k\}$ are Cauchy sequences, we first demonstrate that these sequences are bounded. An effective approach is to show that the Lagrangian function associated with each iteration is bounded. Specifically, based on the update rules for $\dot{\mathbf{L}}_1^k$, $\dot{\mathbf{L}}_2^k$ and $\dot{\mathbf{L}}_3^k$, we have:

$$\begin{aligned}
&\mathcal{L}_{\mu^{k+1}}\left(\dot{\mathbf{Q}}_U^{k+1}, \dot{\mathbf{Q}}_V^{k+1}, \dot{\mathbf{U}}^{k+1}, \dot{\mathbf{V}}^{k+1}, \dot{\mathbf{X}}^{k+1}, \dot{\mathbf{L}}_1^{k+1}, \dot{\mathbf{L}}_2^{k+1}, \dot{\mathbf{L}}_3^{k+1}\right) \\
&= \frac{\lambda}{2}\left(||\dot{\mathbf{Q}}_U^{k+1}||_{\mathbf{w},F}^2 + ||\dot{\mathbf{Q}}_V^{k+1}||_{\mathbf{w},F}^2\right) + \Re\left(\langle \dot{\mathbf{L}}_1^{k+1}, \dot{\mathbf{U}}^{k+1} - \dot{\mathbf{Q}}_U^{k+1}\rangle\right) + \Re\left(\langle \dot{\mathbf{L}}_2^{k+1}, \dot{\mathbf{V}}^{k+1} - \dot{\mathbf{Q}}_V^{k+1}\rangle\right) \\
&+ \Re\left(\langle \dot{\mathbf{L}}_3^{k+1}, \dot{\mathbf{X}}^{k+1} - \dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H\rangle\right) + \frac{\mu^{k+1}}{2}||\dot{\mathbf{U}}^{k+1} - \dot{\mathbf{Q}}_U^{k+1}||_F^2 + \frac{\mu^{k+1}}{2}||\dot{\mathbf{V}}^{k+1} - \dot{\mathbf{Q}}_V^{k+1}||_F^2 \\
&+ \frac{\mu^{k+1}}{2}||\dot{\mathbf{X}}^{k+1} - \dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H||_F^2 + \frac{1}{2}||\mathcal{P}_\Omega(\dot{\mathbf{X}}^{k+1} - \dot{\mathbf{Y}})||_F^2 \\
&= \mathcal{L}_{\mu^k}\left(\dot{\mathbf{Q}}_U^{k+1}, \dot{\mathbf{Q}}_V^{k+1}, \dot{\mathbf{U}}^{k+1}, \dot{\mathbf{V}}^{k+1}, \dot{\mathbf{X}}^{k+1}, \dot{\mathbf{L}}_1^k, \dot{\mathbf{L}}_2^k, \dot{\mathbf{L}}_3^k\right) + \Re\left(\langle \dot{\mathbf{L}}_1^{k+1} - \dot{\mathbf{L}}_1^k, \dot{\mathbf{U}}^{k+1} - \dot{\mathbf{Q}}_U^{k+1}\rangle\right) \\
&+ \Re\left(\langle \dot{\mathbf{L}}_2^{k+1} - \dot{\mathbf{L}}_2^k, \dot{\mathbf{V}}^{k+1} - \dot{\mathbf{Q}}_V^{k+1}\rangle\right) + \Re\left(\langle \dot{\mathbf{L}}_3^{k+1} - \dot{\mathbf{L}}_3^k, \dot{\mathbf{X}}^{k+1} - \dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H\rangle\right) \\
&+ \frac{\mu^{k+1} - \mu^k}{2}\left(||\dot{\mathbf{U}}^{k+1} - \dot{\mathbf{Q}}_U^{k+1}||_F^2 + ||\dot{\mathbf{V}}^{k+1} - \dot{\mathbf{Q}}_V^{k+1}||_F^2 + ||\dot{\mathbf{X}}^{k+1} - \dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H||_F^2\right) \\
&= \mathcal{L}_{\mu^k}\left(\dot{\mathbf{Q}}_U^{k+1}, \dot{\mathbf{Q}}_V^{k+1}, \dot{\mathbf{U}}^{k+1}, \dot{\mathbf{V}}^{k+1}, \dot{\mathbf{X}}^{k+1}, \dot{\mathbf{L}}_1^k, \dot{\mathbf{L}}_2^k, \dot{\mathbf{L}}_3^k\right) + \frac{\mu^{k+1} + \mu^k}{2(\mu^k)^2}\left(||\dot{\mathbf{L}}_1^{k+1} - \dot{\mathbf{L}}_1^k||_F^2 + ||\dot{\mathbf{L}}_2^{k+1} - \dot{\mathbf{L}}_2^k||_F^2 + ||\dot{\mathbf{L}}_3^{k+1} - \dot{\mathbf{L}}_3^k||_F^2\right).
\end{aligned}$$

Let $M_j^2 = \max\left\{||\dot{\mathbf{L}}_j^{k+1} - \dot{\mathbf{L}}_j^k||_F^2, k = 1, 2, \cdots\right\}$ $(j = 1, 2, 3)$, it follows that:

$$\begin{aligned}
&\mathcal{L}_{\mu^{k+1}}\left(\dot{\mathbf{Q}}_U^{k+1}, \dot{\mathbf{Q}}_V^{k+1}, \dot{\mathbf{U}}^{k+1}, \dot{\mathbf{V}}^{k+1}, \dot{\mathbf{X}}^{k+1}, \dot{\mathbf{L}}_1^{k+1}, \dot{\mathbf{L}}_2^{k+1}, \dot{\mathbf{L}}_3^{k+1}\right) \\
&\leq \mathcal{L}_{\mu^k}\left(\dot{\mathbf{Q}}_U^{k+1}, \dot{\mathbf{Q}}_V^{k+1}, \dot{\mathbf{U}}^{k+1}, \dot{\mathbf{V}}^{k+1}, \dot{\mathbf{X}}^{k+1}, \dot{\mathbf{L}}_1^k, \dot{\mathbf{L}}_2^k, \dot{\mathbf{L}}_3^k\right) + \frac{\mu^{k+1} + \mu^k}{2(\mu^k)^2}\left(M_1^2 + M_2^2 + M_3^2\right) \\
&\leq \mathcal{L}_{\mu^0}\left(\dot{\mathbf{Q}}_U^1, \dot{\mathbf{Q}}_V^1, \dot{\mathbf{U}}^1, \dot{\mathbf{V}}^1, \dot{\mathbf{X}}^1, \dot{\mathbf{L}}_1^0, \dot{\mathbf{L}}_2^0, \dot{\mathbf{L}}_3^0\right) + \left(M_1^2 + M_2^2 + M_3^2\right)\sum_{i=0}^{k}\frac{1+\rho}{2\mu^0\rho^i} \\
&\leq \mathcal{L}_{\mu^0}\left(\dot{\mathbf{Q}}_U^1, \dot{\mathbf{Q}}_V^1, \dot{\mathbf{U}}^1, \dot{\mathbf{V}}^1, \dot{\mathbf{X}}^1, \dot{\mathbf{L}}_1^0, \dot{\mathbf{L}}_2^0, \dot{\mathbf{L}}_3^0\right) + \frac{\left(M_1^2 + M_2^2 + M_3^2\right)}{\mu^0}\sum_{i=0}^{k}\frac{1}{\rho^{i-1}} < +\infty.
\end{aligned}$$

It is noted that the second inequality holds universally, since the globally optimal solutions for $\{\dot{\mathbf{X}}^k\}, \{\dot{\mathbf{Q}}_U^k\}$ and $\{\dot{\mathbf{Q}}_V^k\}$ are obtained in their respective subproblems. This ensures that the sequence of the Lagrangian function is bounded. Furthermore, we proceed to show that the sequences $\{\dot{\mathbf{X}}^k\}$, $\{\dot{\mathbf{Q}}_U^k\}$ and $\{\dot{\mathbf{Q}}_V^k\}$

are bounded.

$$
\begin{aligned}
&\frac{1}{2}||\mathcal{P}_\Omega(\dot{\mathbf{X}}^{k+1}-\dot{\mathbf{Y}})||_F^2+\frac{\lambda}{2}\left(||\dot{\mathbf{Q}}_U^{k+1}||_{\mathbf{w},F}^2+||\dot{\mathbf{Q}}_V^{k+1}||_{\mathbf{w},F}^2\right)\\
&=\mathcal{L}_{\mu^k}\left(\dot{\mathbf{Q}}_U^{k+1},\dot{\mathbf{Q}}_V^{k+1},\dot{\mathbf{U}}^{k+1},\dot{\mathbf{V}}^{k+1},\dot{\mathbf{X}}^{k+1},\dot{\mathbf{L}}_1^k,\dot{\mathbf{L}}_2^k,\dot{\mathbf{L}}_3^k\right)-\Re\left(\langle\dot{\mathbf{L}}_1^k,\dot{\mathbf{U}}^{k+1}-\dot{\mathbf{Q}}_U^{k+1}\rangle\right)-\Re\left(\langle\dot{\mathbf{L}}_2^k,\dot{\mathbf{V}}^{k+1}-\dot{\mathbf{Q}}_V^{k+1}\rangle\right)\\
&-\Re\left(\langle\dot{\mathbf{L}}_3^k,\dot{\mathbf{X}}^{k+1}-\dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H\rangle\right)-\frac{\mu^k}{2}||\dot{\mathbf{U}}^{k+1}-\dot{\mathbf{Q}}_U^{k+1}||_F^2-\frac{\mu^k}{2}||\dot{\mathbf{V}}^{k+1}-\dot{\mathbf{Q}}_V^{k+1}||_F^2-\frac{\mu^k}{2}||\dot{\mathbf{X}}^{k+1}-\dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H||_F^2\\
&=\mathcal{L}_{\mu^k}\left(\dot{\mathbf{Q}}_U^{k+1},\dot{\mathbf{Q}}_V^{k+1},\dot{\mathbf{U}}^{k+1},\dot{\mathbf{V}}^{k+1},\dot{\mathbf{X}}^{k+1},\dot{\mathbf{L}}_1^k,\dot{\mathbf{L}}_2^k,\dot{\mathbf{L}}_3^k\right)\\
&+\frac{1}{2\mu^k}\left(||\dot{\mathbf{L}}_1^k||_F^2-||\dot{\mathbf{L}}_1^{k+1}||_F^2+||\dot{\mathbf{L}}_2^k||_F^2-||\dot{\mathbf{L}}_2^{k+1}||_F^2+||\dot{\mathbf{L}}_3^k||_F^2-||\dot{\mathbf{L}}_3^{k+1}||_F^2\right).
\end{aligned}
$$

Note that $||\mathcal{P}_\Omega(\dot{\mathbf{X}}^{k+1}-\dot{\mathbf{Y}})||_F^2$ and the reweighted Frobenius norm are nonnegative. Therefore, the sequences $\{\dot{\mathbf{X}}^k\},\{\dot{\mathbf{Q}}_U^k\}$ and $\{\dot{\mathbf{Q}}_V^k\}$ are bounded. Furthermore, based on (E.1), it follows that sequences $\{\dot{\mathbf{U}}^k\}$ and $\{\dot{\mathbf{V}}^k\}$ are also bounded.

From (13), by applying quaternion matrix derivatives and setting them to zero, we substitute

$$
\dot{\mathbf{X}}^k=\dot{\mathbf{U}}^k(\dot{\mathbf{V}}^k)^H+\frac{\dot{\mathbf{L}}_3^k}{\mu^{k-1}}-\frac{\dot{\mathbf{L}}_3^{k-1}}{\mu^{k-1}}
$$

to derive the following computational procedure:

$$
\begin{aligned}
&\left(\dot{\mathbf{U}}^{k+1}-\dot{\mathbf{Q}}_U^k+\frac{\dot{\mathbf{L}}_1^k}{\mu^k}\right)+\left(\dot{\mathbf{U}}^{k+1}\left(\dot{\mathbf{V}}^k\right)^H-\dot{\mathbf{X}}^k-\frac{\dot{\mathbf{L}}_3^k}{\mu^k}\right)\dot{\mathbf{V}}^k\\
&=\left(\dot{\mathbf{U}}^{k+1}-\dot{\mathbf{U}}^k+\dot{\mathbf{U}}^k-\dot{\mathbf{Q}}_U^k+\frac{\dot{\mathbf{L}}_1^k}{\mu^k}\right)+\left(\dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^k)^H-\dot{\mathbf{U}}^k(\dot{\mathbf{V}}^k)^H-\frac{\dot{\mathbf{L}}_3^k}{\mu^{k-1}}+\frac{\dot{\mathbf{L}}_3^{k-1}}{\mu^{k-1}}-\frac{\dot{\mathbf{L}}_3^k}{\mu^k}\right)\dot{\mathbf{V}}^k\\
&=(\dot{\mathbf{U}}^{k+1}-\dot{\mathbf{U}}^k)\left(\dot{\mathbf{I}}+(\dot{\mathbf{V}}^k)^H\dot{\mathbf{V}}^k\right)+\frac{\dot{\mathbf{L}}_1^k-\dot{\mathbf{L}}_1^{k-1}}{\mu^{k-1}}+\frac{\dot{\mathbf{L}}_1^k}{\mu^k}-\left(\frac{\dot{\mathbf{L}}_3^k-\dot{\mathbf{L}}_3^{k-1}}{\mu^{k-1}}+\frac{\dot{\mathbf{L}}_3^k}{\mu^k}\right)\dot{\mathbf{V}}^k=\dot{\mathbf{0}},
\end{aligned}
$$

and

$$
\begin{aligned}
&\left(\dot{\mathbf{V}}^{k+1}-\dot{\mathbf{Q}}_V^k+\frac{\dot{\mathbf{L}}_2^k}{\mu^k}\right)+\left(\dot{\mathbf{V}}^{k+1}(\dot{\mathbf{U}}^{k+1})^H-(\dot{\mathbf{X}}^k)^H-\frac{(\dot{\mathbf{L}}_3^k)^H}{\mu^k}\right)\dot{\mathbf{U}}^{k+1}\\
&=\left(\dot{\mathbf{V}}^{k+1}-\dot{\mathbf{V}}^k+\dot{\mathbf{V}}^k-\dot{\mathbf{Q}}_V^k+\frac{\dot{\mathbf{L}}_2^k}{\mu^k}\right)+\left(\dot{\mathbf{V}}^{k+1}(\dot{\mathbf{U}}^{k+1})^H-\dot{\mathbf{V}}^k(\dot{\mathbf{U}}^k)^H-\frac{(\dot{\mathbf{L}}_3^k)^H}{\mu^{k-1}}+\frac{(\dot{\mathbf{L}}_3^{k-1})^H}{\mu^{k-1}}-\frac{(\dot{\mathbf{L}}_3^k)^H}{\mu^k}\right)\dot{\mathbf{U}}^{k+1}\\
&=\left(\dot{\mathbf{V}}^{k+1}-\dot{\mathbf{V}}^k+\dot{\mathbf{V}}^k-\dot{\mathbf{Q}}_V^k+\frac{\dot{\mathbf{L}}_2^k}{\mu^k}\right)+(\dot{\mathbf{V}}^{k+1}(\dot{\mathbf{U}}^{k+1})^H\\
&-\dot{\mathbf{V}}^k(\dot{\mathbf{U}}^{k+1})^H+\dot{\mathbf{V}}^k(\dot{\mathbf{U}}^{k+1})^H-\dot{\mathbf{V}}^k(\dot{\mathbf{U}}^k)^H-\frac{(\dot{\mathbf{L}}_3^k)^H}{\mu^{k-1}}+\frac{(\dot{\mathbf{L}}_3^{k-1})^H}{\mu^{k-1}}-\frac{(\dot{\mathbf{L}}_3^k)^H}{\mu^k})\dot{\mathbf{U}}^{k+1}\\
&=(\dot{\mathbf{V}}^{k+1}-\dot{\mathbf{V}}^k)\left(\dot{\mathbf{I}}+(\dot{\mathbf{U}}^{k+1})^H\dot{\mathbf{U}}^{k+1}\right)+\frac{\dot{\mathbf{L}}_2^k-\dot{\mathbf{L}}_2^{k-1}}{\mu^{k-1}}+\frac{\dot{\mathbf{L}}_2^k}{\mu^k}+\dot{\mathbf{V}}^k(\dot{\mathbf{U}}^{k+1}\\
&-\dot{\mathbf{U}}^k)^H\dot{\mathbf{U}}^{k+1}-\left(\frac{(\dot{\mathbf{L}}_3^k-\dot{\mathbf{L}}_3^{k-1})^H}{\mu^{k-1}}+\frac{(\dot{\mathbf{L}}_3^k)^H}{\mu^k}\right)\dot{\mathbf{U}}^{k+1}=\dot{\mathbf{0}}.
\end{aligned}
$$

To proceed, we isolate the differences $\dot{\mathbf{U}}^{k+1} - \dot{\mathbf{U}}^k$ and $\dot{\mathbf{V}}^{k+1} - \dot{\mathbf{V}}^k$ on one side of the equation, which yields:

$$\begin{aligned}
&\dot{\mathbf{U}}^{k+1} - \dot{\mathbf{U}}^k \\
&= \left(\frac{\dot{\mathbf{L}}_1^{k-1} - \dot{\mathbf{L}}_1^k}{\mu^{k-1}} - \frac{\dot{\mathbf{L}}_1^k}{\mu^k} + \left(\frac{\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1}}{\mu^{k-1}} + \frac{\dot{\mathbf{L}}_3^k}{\mu^k} \right) \dot{\mathbf{V}}^k \right) \left(\dot{\mathbf{I}} + \left(\dot{\mathbf{V}}^k \right)^H \dot{\mathbf{V}}^k \right)^{-1} \\
&= \frac{1}{\mu^k} \left(\rho(\dot{\mathbf{L}}_1^{k-1} - \dot{\mathbf{L}}_1^k) - \dot{\mathbf{L}}_1^k + \left(\rho(\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1}) + \dot{\mathbf{L}}_3^k \right) \dot{\mathbf{V}}^k \right) \left(\dot{\mathbf{I}} + (\dot{\mathbf{V}}^k)^H \dot{\mathbf{V}}^k \right)^{-1} = \frac{1}{\mu^k} \dot{\mathbf{C}}_1^k,
\end{aligned}$$

and

$$\begin{aligned}
&\dot{\mathbf{V}}^{k+1} - \dot{\mathbf{V}}^k \\
&= \left(\frac{\dot{\mathbf{L}}_2^{k-1} - \dot{\mathbf{L}}_2^k}{\mu^{k-1}} - \frac{\dot{\mathbf{L}}_2^k}{\mu^k} + \dot{\mathbf{V}}^k (\dot{\mathbf{U}}^k - \dot{\mathbf{U}}^{k+1})^H \dot{\mathbf{U}}^{k+1} + \left(\frac{(\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1})^H}{\mu^{k-1}} + \frac{(\dot{\mathbf{L}}_3^k)^H}{\mu^k} \right) \dot{\mathbf{U}}^{k+1} \right) \left(\dot{\mathbf{I}} + (\dot{\mathbf{U}}^{k+1})^H \dot{\mathbf{U}}^{k+1} \right)^{-1} \\
&= \frac{1}{\mu^k} \left(\rho(\dot{\mathbf{L}}_2^{k-1} - \dot{\mathbf{L}}_2^k) - \dot{\mathbf{L}}_2^k + \mu^k \dot{\mathbf{V}}^k (\dot{\mathbf{U}}^k - \dot{\mathbf{U}}^{k+1})^H \dot{\mathbf{U}}^{k+1} + \left(\rho(\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1})^H + (\dot{\mathbf{L}}_3^k)^H \right) \dot{\mathbf{U}}^{k+1} \right) \left(\dot{\mathbf{I}} + (\dot{\mathbf{U}}^{k+1})^H \dot{\mathbf{U}}^{k+1} \right)^{-1} \\
&= \frac{1}{\mu^k} \left(\rho(\dot{\mathbf{L}}_2^{k-1} - \dot{\mathbf{L}}_2^k) - \dot{\mathbf{L}}_2^k - \dot{\mathbf{V}}^k (\dot{\mathbf{C}}_1^k)^H \dot{\mathbf{U}}^{k+1} + \left(\rho(\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1})^H + (\dot{\mathbf{L}}_3^k)^H \right) \dot{\mathbf{U}}^{k+1} \right) \left(\dot{\mathbf{I}} + (\dot{\mathbf{U}}^{k+1})^H \dot{\mathbf{U}}^{k+1} \right)^{-1} \\
&= \frac{1}{\mu^k} \dot{\mathbf{C}}_2^k,
\end{aligned}$$

where $\dot{\mathbf{C}}_1^k = \left(\rho(\dot{\mathbf{L}}_1^{k-1} - \dot{\mathbf{L}}_1^k) - \dot{\mathbf{L}}_1^k + \left(\rho(\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1}) + \dot{\mathbf{L}}_3^k \right) \dot{\mathbf{V}}^k \right) \left(\dot{\mathbf{I}} + (\dot{\mathbf{V}}^k)^H \dot{\mathbf{V}}^k \right)^{-1}$, $\dot{\mathbf{C}}_2^k = \left(\rho(\dot{\mathbf{L}}_2^{k-1} - \dot{\mathbf{L}}_2^k) - \dot{\mathbf{L}}_2^k - \dot{\mathbf{V}}^k (\dot{\mathbf{C}}_1^k)^H \dot{\mathbf{U}}^{k+1} + \left(\rho(\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1})^H + (\dot{\mathbf{L}}_3^k)^H \right) \dot{\mathbf{U}}^{k+1} \right) \left(\dot{\mathbf{I}} + (\dot{\mathbf{U}}^{k+1})^H \dot{\mathbf{U}}^{k+1} \right)^{-1}$. Therefore, the sequences $\{\dot{\mathbf{U}}^k\}$ and $\{\dot{\mathbf{V}}^k\}$ are Cauchy sequences.

In the following, we demonstrate that the sequence $\{\dot{\mathbf{X}}^k\}$ is also a Cauchy sequence. Based on the update step for the sequence $\{\dot{\mathbf{L}}_3^k\}$, we have:

$$\dot{\mathbf{X}}^{k+1} = \frac{\dot{\mathbf{L}}_3^{k+1} - \dot{\mathbf{L}}_3^k}{\mu^k} + \dot{\mathbf{U}}^{k+1} (\dot{\mathbf{V}}^{k+1})^H,$$

and

$$\begin{aligned}
&||\dot{\mathbf{X}}^{k+1} - \dot{\mathbf{X}}^k||_F \\
&= ||\mathcal{P}_\Omega \left(\frac{\mu^k \dot{\mathbf{U}}^{k+1} (\dot{\mathbf{V}}^{k+1})^H - \dot{\mathbf{L}}_3^k + \dot{\mathbf{Y}}}{1 + \mu^k} \right) + \mathcal{P}_{\Omega^c} \left(\dot{\mathbf{U}}^{k+1} (\dot{\mathbf{V}}^{k+1})^H - \frac{\dot{\mathbf{L}}_3^k}{\mu^k} \right) - \frac{\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1}}{\mu^{k-1}} - \dot{\mathbf{U}}^k (\dot{\mathbf{V}}^k)^H ||_F \\
&= ||\mathcal{P}_\Omega \left(\frac{\dot{\mathbf{Y}} - \dot{\mathbf{L}}_3^k - \dot{\mathbf{U}}^{k+1} (\dot{\mathbf{V}}^{k+1})^H}{1 + \mu^k} \right) - \mathcal{P}_{\Omega^c} \left(\frac{\dot{\mathbf{L}}_3^k}{\mu^k} \right) - \frac{\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1}}{\mu^{k-1}} + (\dot{\mathbf{U}}^{k+1} - \dot{\mathbf{U}}^k)(\dot{\mathbf{V}}^{k+1})^H + \dot{\mathbf{U}}^k (\dot{\mathbf{V}}^{k+1} - \dot{\mathbf{V}}^k)^H ||_F \\
&= ||\mathcal{P}_\Omega \left(\frac{\dot{\mathbf{Y}} - \dot{\mathbf{L}}_3^k - \dot{\mathbf{U}}^{k+1} (\dot{\mathbf{V}}^{k+1})^H}{1 + \mu^k} \right) - \mathcal{P}_{\Omega^c} \left(\frac{\dot{\mathbf{L}}_3^k}{\mu^k} \right) - \frac{\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1}}{\mu^{k-1}} + \frac{\dot{\mathbf{C}}_1^k}{\mu^k} (\dot{\mathbf{V}}^{k+1})^H + \frac{\dot{\mathbf{U}}^k}{\mu^k} (\dot{\mathbf{C}}_2^k)^H ||_F \\
&= \frac{1}{\mu^k} ||\dot{\mathbf{C}}_3^k||_F,
\end{aligned}$$

where $\dot{\mathbf{C}}_3^k = \mathcal{P}_\Omega \left(\frac{\mu^k}{1+\mu^k} \left(\dot{\mathbf{Y}} - \dot{\mathbf{L}}_3^k - \dot{\mathbf{U}}^{k+1} (\dot{\mathbf{V}}^{k+1})^H \right) \right) - \mathcal{P}_{\Omega^c}(\dot{\mathbf{L}}_3^k) - \rho(\dot{\mathbf{L}}_3^k - \dot{\mathbf{L}}_3^{k-1}) + \dot{\mathbf{C}}_1^k (\dot{\mathbf{V}}^{k+1})^H + \dot{\mathbf{U}}^k (\dot{\mathbf{C}}_2^k)^H$. Thus, $\{\dot{\mathbf{X}}^k\}$ is a Cauchy sequence.

2. By applying the first-order optimization condition to equations (13), (17) and (20), we have:

$$\left\{\begin{array}{l} \dot{\mathbf{0}} \in \lambda\partial||\dot{\mathbf{Q}}_U^{k+1}||_{\mathbf{w},F} + \mu^k \left(\dot{\mathbf{Q}}_U^{k+1} - \dot{\mathbf{U}}^{k+1} - \frac{\dot{\mathbf{L}}_1^k}{\mu^k}\right), \\ \dot{\mathbf{0}} \in \lambda\partial||\dot{\mathbf{Q}}_V^{k+1}||_{\mathbf{w},F} + \mu^k \left(\dot{\mathbf{Q}}_V^{k+1} - \dot{\mathbf{V}}^{k+1} - \frac{\dot{\mathbf{L}}_2^k}{\mu^k}\right), \\ \mathcal{P}_\Omega\left(\dot{\mathbf{X}}^{k+1} - \frac{\mu^k\dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H - \dot{\mathbf{L}}_3^k + \dot{\mathbf{Y}}}{1+\mu^k}\right) + \mathcal{P}_{\Omega^c}\left(\dot{\mathbf{X}}^{k+1} - \dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H + \frac{\dot{\mathbf{L}}_3^k}{\mu^k}\right) = \dot{\mathbf{0}}. \end{array}\right.$$

Let $\dot{\mathbf{U}}^\star$, $\dot{\mathbf{V}}^\star$, $\dot{\mathbf{Q}}_U^\star$, $\dot{\mathbf{Q}}_V^\star$ and $\dot{\mathbf{X}}^\star$ denote the accumulation points of the sequences $\{\dot{\mathbf{U}}^k\}$, $\{\dot{\mathbf{V}}^k\}$, $\{\dot{\mathbf{Q}}_U^k\}$, $\{\dot{\mathbf{Q}}_V^k\}$ and $\{\dot{\mathbf{X}}^k\}$, respectively. It follows that $\dot{\mathbf{U}}^\star = \dot{\mathbf{Q}}_U^\star$, $\dot{\mathbf{V}}^\star = \dot{\mathbf{Q}}_V^\star$ and $\dot{\mathbf{X}}^\star = \dot{\mathbf{U}}^\star(\dot{\mathbf{V}}^\star)^H$. Therefore, as $k \to +\infty$, it obtains:

$$\left\{\begin{array}{l} \dot{\mathbf{0}} \in \lambda\partial||\dot{\mathbf{Q}}_U^\star||_{\mathbf{w},F} - \dot{\mathbf{L}}_1^\star, \\ \dot{\mathbf{0}} \in \lambda\partial||\dot{\mathbf{Q}}_V^\star||_{\mathbf{w},F} - \dot{\mathbf{L}}_2^\star, \\ \mathcal{P}_\Omega(\dot{\mathbf{X}}^\star) - \dot{\mathbf{Y}} + \mathcal{P}_\Omega(\dot{\mathbf{L}}_3^\star) = \dot{\mathbf{0}}, \\ \mathcal{P}_{\Omega^c}(\dot{\mathbf{L}}_3^\star) = \dot{\mathbf{0}}. \end{array}\right.$$

Hence, it concludes that any accumulation points $\{(\dot{\mathbf{U}}^\star, \dot{\mathbf{V}}^\star, \dot{\mathbf{Q}}_U^\star, \dot{\mathbf{A}}_V^\star, \dot{\mathbf{X}}^\star)\}$ of the sequence $\{(\dot{\mathbf{U}}^k, \dot{\mathbf{V}}^k, \dot{\mathbf{Q}}_U^k, \dot{\mathbf{Q}}_V^k, \dot{\mathbf{X}}^k)\}$ generated by QRLMF satisfies the KKT conditions.

Appendix F. The proof of Theorem 5

1. As in the proof of Theorem 4, it can also be verified that the sequences $\{\dot{\mathbf{Q}}_U^k\}$, $\{\dot{\mathbf{Q}}_V^k\}$, $\{\dot{\mathbf{U}}^k\}$ and $\{\dot{\mathbf{V}}^k\}$ are Cauchy sequences. In the following, we proceed to show that the sequences $\{\dot{\mathbf{X}}^k\}$ and $\{\mathcal{S}^k\}$ also are Cauchy sequences.

By analyzing equation (30), from which we can derive that:

$$\left\{\begin{array}{l} \dot{\mathbf{X}}^{k+1} = \frac{\dot{\mathbf{L}}_3^{k+1} - \dot{\mathbf{L}}_3^k}{\mu^k} + \dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H, \\ \Gamma^{-1}(\dot{\mathbf{X}}^{k+1}) = \frac{\mathcal{T}^{k+1} - \mathcal{T}^k}{\mu^k} + \mathcal{S}^{k+1}. \end{array}\right.$$

Convert $\Gamma^{-1}(\dot{\mathbf{X}}^{k+1})$ into quaternion form. Note that $\dot{\mathbf{X}}^k$ and $\Gamma^{-1}(\dot{\mathbf{X}}^k)$ represent the same color image but in different formats. Consequently, we have:

$$\begin{aligned} \dot{\mathbf{X}}^{k+1} &= \frac{\dot{\mathbf{X}}^{k+1} + \Gamma\left(\Gamma^{-1}(\dot{\mathbf{X}}^{k+1})\right)}{2} \\ &= \frac{\frac{\dot{\mathbf{L}}_3^{k+1} - \dot{\mathbf{L}}_3^k}{\mu^k} + \dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H + \Gamma\left(\frac{\mathcal{T}^{k+1} - \mathcal{T}^k}{\mu^k}\right) + \Gamma(\mathcal{S}^{k+1})}{2}. \end{aligned}$$

Furthermore,

$$
\begin{aligned}
&||\dot{\mathbf{X}}^{k+1}-\dot{\mathbf{X}}^{k}||_F\\
&=||\mathcal{P}_\Omega\left(\frac{\mu^k\dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H+\mu^k\Gamma(\mathcal{S}^k)-\dot{\mathbf{L}}_3^k-\Gamma(\mathcal{T}^k)+\dot{\mathbf{Y}}}{1+2\mu^k}\right)+\mathcal{P}_{\Omega^c}\left(\frac{1}{2}\left(\dot{\mathbf{U}}^{k+1}(\dot{\mathbf{V}}^{k+1})^H+\Gamma(\mathcal{S}^k)-\frac{\dot{\mathbf{L}}_3^k}{\mu^k}-\Gamma\left(\frac{\mathcal{T}^k}{\mu^k}\right)\right)\right)\\
&-\frac{1}{2}\left(\dot{\mathbf{U}}^{k}(\dot{\mathbf{V}}^{k})^H+\Gamma(\mathcal{S}^k)+\frac{\dot{\mathbf{L}}_3^k-\dot{\mathbf{L}}_3^{k-1}}{\mu^{k-1}}+\Gamma\left(\frac{\mathcal{T}^k-\mathcal{T}^{k-1}}{\mu^{k-1}}\right)\right)||_F\\
&=||\mathcal{P}_\Omega\left(\frac{2\dot{\mathbf{Y}}-\Gamma(\mathcal{S}^k)-2\dot{\mathbf{L}}_3^k-2\Gamma(\mathcal{T}^k)-\dot{\mathbf{U}}^{k}(\dot{\mathbf{V}}^{k})^H}{2(1+2\mu^k)}-\frac{\dot{\mathbf{C}}_1^k(\dot{\mathbf{V}}^{k+1})^H-\dot{\mathbf{U}}^{k}(\dot{\mathbf{C}}_2^k)^H}{(1+2\mu^k)}\right)+\frac{1}{2}\mathcal{P}_{\Omega^c}\left(\frac{1}{\mu^k}\left(\dot{\mathbf{C}}_1^k(\dot{\mathbf{V}}^{k+1})^H-\dot{\mathbf{U}}^{k}(\dot{\mathbf{C}}_2^k)^H\right)\right)\\
&-\frac{1}{2}\mathcal{P}_{\Omega^c}\left(\frac{\dot{\mathbf{L}}_3^k}{\mu^k}+\Gamma\left(\frac{\mathcal{T}^k}{\mu^k}\right)\right)-\frac{\dot{\mathbf{L}}_3^k-\dot{\mathbf{L}}_3^{k-1}}{2\mu^{k-1}}-\Gamma\left(\frac{\mathcal{T}^k-\mathcal{T}^{k-1}}{2\mu^{k-1}}\right)||_F=\frac{1}{\mu^k}||\dot{\mathbf{C}}_4^k||_F,
\end{aligned}
$$

where $\dot{\mathbf{C}}_4^k=\mathcal{P}_\Omega\left(\frac{\mu^k}{2(1+2\mu^k)}\left(2\dot{\mathbf{Y}}-\Gamma(\mathcal{S}^k)-2\dot{\mathbf{L}}_3^k-2\Gamma(\mathcal{T}^k)-\dot{\mathbf{U}}^{k}(\dot{\mathbf{V}}^{k})^H-2\dot{\mathbf{C}}_1^k(\dot{\mathbf{V}}^{k+1})^H-2\dot{\mathbf{U}}^{k}(\dot{\mathbf{C}}_2^k)^H\right)\right)$ $+\frac{1}{2}\mathcal{P}_{\Omega^c}\left(\dot{\mathbf{C}}_1^k(\dot{\mathbf{V}}^{k+1})^H+\dot{\mathbf{U}}^{k}(\dot{\mathbf{C}}_2^k)^H\right)-\frac{1}{2}\mathcal{P}_{\Omega^c}\left(\dot{\mathbf{L}}_3^k+\Gamma(\mathcal{T}^k)\right)-\frac{\rho(\dot{\mathbf{L}}_3^k-\dot{\mathbf{L}}_3^{k-1})}{2}-\rho\Gamma\left(\frac{\mathcal{T}^k-\mathcal{T}^{k-1}}{2}\right)$. Therefore, it follows that $\{\dot{\mathbf{X}}^k\}$ is a Cauchy sequence.

Based on (30), we have:

$$
\mathcal{T}^{k+1}=\mathcal{T}^k+\mu^k\left(\Gamma^{-1}(\dot{\mathbf{X}}^{k+1})-\mathcal{S}^{k+1}\right),
$$

it is easy to obtain:

$$
\mathcal{S}^{k+1}=\Gamma^{-1}(\dot{\mathbf{X}}^{k+1})+\frac{\mathcal{T}^k}{\mu^k}-\frac{\mathcal{T}^{k+1}}{\mu^k}.
$$

Furthermore, we prove that the sequence $\{\mathcal{S}^k\}$ is a Cauchy sequence.

$$
\begin{aligned}
||\mathcal{S}^{k+1}-\mathcal{S}^k||_F&=||\Gamma^{-1}(\dot{\mathbf{X}}^{k+1})+\frac{\mathcal{T}^k}{\mu^k}-\frac{\mathcal{T}^{k+1}}{\mu^k}-\mathcal{S}^k||_F\\
&=||\Gamma^{-1}(\dot{\mathbf{X}}^{k})-\mathcal{S}^k+\Gamma^{-1}(\dot{\mathbf{X}}^{k+1})-\Gamma^{-1}(\dot{\mathbf{X}}^{k})+\frac{\mathcal{T}^k-\mathcal{T}^{k+1}}{\mu^k}||_F\\
&=||\frac{\mathcal{T}^k-\mathcal{T}^{k-1}}{\mu^{k-1}}+\frac{\Gamma^{-1}(\dot{\mathbf{C}}_4^k)}{\mu^k}+\frac{\mathcal{T}^k-\mathcal{T}^{k+1}}{\mu^k}||_F=\frac{1}{\mu^k}||\mathcal{C}_5^k||_F,
\end{aligned}
$$

where $\mathcal{C}_5^k=\rho(\mathcal{T}^k-\mathcal{T}^{k-1})+\Gamma^{-1}(\dot{\mathbf{C}}_4^k)+\mathcal{T}^k-\mathcal{T}^{k+1}$. Therefore, it follows that $\{\mathcal{S}^k\}$ is a Cauchy sequence.

2. By applying the first-order optimization condition to equations (17), (28) and (29), and let $\dot{\mathbf{U}}^\star$, $\dot{\mathbf{V}}^\star$, $\dot{\mathbf{Q}}_U^\star$, $\dot{\mathbf{Q}}_V^\star$, $\dot{\mathbf{X}}^\star$ and $\mathcal{S}^\star$ denote the accumulation points of the sequences $\{\dot{\mathbf{U}}^k\}$, $\{\dot{\mathbf{V}}^k\}$, $\{\dot{\mathbf{Q}}_U^k\}$, $\{\dot{\mathbf{Q}}_V^k\}$, $\{\dot{\mathbf{X}}^k\}$ and $\{\mathcal{S}^k\}$, respectively. We observe that since these sequences are Cauchy sequences, it follows that $\dot{\mathbf{U}}^\star=\dot{\mathbf{Q}}_U^\star$, $\dot{\mathbf{V}}^\star=\dot{\mathbf{Q}}_V^\star$, $\dot{\mathbf{X}}^\star=\dot{\mathbf{U}}^\star(\dot{\mathbf{V}}^\star)^H$ and $\Gamma^{-1}(\dot{\mathbf{X}}^\star)=\mathcal{S}^\star$. Thus, as $k\to+\infty$, we have:

$$
\begin{cases}
\dot{\mathbf{0}} \in \lambda \partial ||\dot{\mathbf{Q}}_U^\star||_{\mathbf{w},F} - \dot{\mathbf{L}}_1^\star, \\
\dot{\mathbf{0}} \in \lambda \partial ||\dot{\mathbf{Q}}_V^\star||_{\mathbf{w},F} - \dot{\mathbf{L}}_2^\star, \\
\mathbf{0} \in \alpha \partial \Phi_{\mathsf{pnp}}(\mathcal{S}^\star) - \mathcal{T}^\star, \\
\mathcal{P}_\Omega(\dot{\mathbf{X}}^\star) - \dot{\mathbf{Y}} + \mathcal{P}_\Omega(\dot{\mathbf{L}}_3^\star) + \mathcal{P}_\Omega(\Gamma(\mathcal{T}^\star)) = \dot{\mathbf{0}}, \\
\mathcal{P}_{\Omega^c}(\dot{\mathbf{L}}_3^\star) + \mathcal{P}_{\Omega^c}(\Gamma(\mathcal{T}^\star)) = \dot{\mathbf{0}}.
\end{cases}
$$

Therefore, any accumulation points $\{(\dot{\mathbf{U}}^\star,\ \dot{\mathbf{V}}^\star,\ \dot{\mathbf{Q}}_U^\star,\ \dot{\mathbf{A}}_V^\star,\ \dot{\mathbf{X}}^\star,\ \mathcal{S}^\star)\}$ of the sequence $\{(\dot{\mathbf{U}}^k,\ \dot{\mathbf{V}}^k,\ \dot{\mathbf{Q}}_U^k,\ \dot{\mathbf{Q}}_V^k,\ \dot{\mathbf{X}}^k,\ \mathcal{S}^k)\}$ generated by DeepQRLMF satisfies the KKT conditions.

References

[1] W. R. Hamilton, *Elements of Quaternions.* Harlow, U. K.: Longmans, 1866.

[2] R. D. Schafer, "On the algebras formed by the Cayley–Dickson process", *Amer. J. Math.,* vol. 76, no. 2, pp. 435–446, 1954.

[3] Y. Chen, X. Xiao, and Y. Zhou, "Low-rank quaternion approximation for color image processing", *IEEE Trans. Image Process.,* vol. 29, pp. 1426-1439, 2020.

[4] F. Zhang, "Quaternions and matrices of quaternions", *Linear Algebra Appl.,* vol. 251, pp. 21-57, 1997.

[5] J. Miao, and K. I. Kou, "Quaternion-based bilinear factor matrix norm minimization for color image inpainting", *IEEE Trans. Signal Process.,* vol. 68, pp. 5617-5631, 2020.

[6] R. A. Horn, and C. R. Johnson, *Topics in Matrix Analysis.* Cambridge, U. K.: Cambridge Univ. Press, 1991.