APPENDIX

The appendices are organized as follows:

- In Appendix A, we provide a detailed exposition of quaternions and their role in representing rotations, along with several fundamental definitions, notations, and lemmas essential for subsequent analysis.
- In Appendix B, we characterize the recovery accuracy of the solution of (Q-QP) with respect to the ground truth $\tilde{q}$.
- In Appendix C, we establish the tightness of semidefinite relaxation, which ensures that the rank-one solution of the convex problem (Q-SDP) yields a global optimum for the nonconvex problem (Q-QP).

A. NOTATION AND PRELIMINARIES

A.1. Basic Notations

The fields of real numbers, quaternion numbers, and unit quaternion numbers are denoted by $\mathbb{R}$, $\mathbb{Q}$, and $\mathbb{U}$, respectively. Throughout this paper, scalars, vectors, matrices, and quaternions are denoted by lowercase letters (e.g., x), boldface lowercase letters (e.g., $\boldsymbol{x}$), boldface capital letters (e.g., X), and lowercase letters with tilde (e.g., $\tilde{x}$), respectively. The special orthogonal group $SO(3)$ is the set of three-dimensional rotations that is formally defined by $SO(3) := \{R \in \mathbb{R}^{3\times 3} \mid R^\top R = I_3, \det(R) = 1\}$. We denote by $\mathbb{S}^{n\times n}$ the set of all symmetric matrices. $\mathbb{1}$ is a vector of all-ones. The notation $\|\cdot\|_2$ denotes the ℓ_2-norm of vectors or the spectral norm of matrices.

A.2. Quaternion and rotation

A quaternion number $\tilde{q} \in \mathbb{Q}$, proposed by Hamilton, has the form $\tilde{q} = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, where $q_0, q_1, q_2, q_3 \in \mathbb{R}$ and $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are three imaginary units. We may also write $\tilde{q} = [q_0, q_1, q_2, q_3] = [q_0, \boldsymbol{q}] \in \mathbb{R}^4$ as the vector representation where $\boldsymbol{q} = [q_1, q_2, q_3] \in \mathbb{R}^3$ for convenience. We note that we also regard the above representation as a column vector and its transpose $[q_0, \boldsymbol{q}]^\top$ a row vector. The sum of $\tilde{p}$ and $\tilde{q}$ is defined as $\tilde{p} + \tilde{q} = [p_0 + q_0, \boldsymbol{p} + \boldsymbol{q}]$. The product of $\tilde{p}$ and $\tilde{q}$ is defined by

$$\tilde{p}\tilde{q} = [p_0 q_0 - \boldsymbol{p}\cdot\boldsymbol{q}, p_0\boldsymbol{q} + q_0\boldsymbol{p} + \boldsymbol{p}\times\boldsymbol{q}],$$

where $\boldsymbol{p}\cdot\boldsymbol{q}$ is the dot product, and $\boldsymbol{p}\times\boldsymbol{q}$ is the cross product of $\boldsymbol{p}$ and $\boldsymbol{q}$. Thus, in general, $\tilde{p}\tilde{q} \neq \tilde{q}\tilde{p}$, and we have $\tilde{p}\tilde{q} = \tilde{q}\tilde{p}$ if and only if $\boldsymbol{p}\times\boldsymbol{q} = \vec{0}$, i.e., either $\boldsymbol{p} = \vec{0}$ or $\boldsymbol{q} = \vec{0}$, or $\boldsymbol{p} = \alpha\boldsymbol{q}$ for several real number α. The multiplication of quaternions is associative and distributive over vector addition, but is not commutative.

The conjugate of $\tilde{q}$ is the quaternion $\tilde{q}^* = q_0 - q_1\mathbf{i} - q_2\mathbf{j} - q_3\mathbf{k}$. Then, $(\tilde{p}\tilde{q})^* = \tilde{q}^*\tilde{p}^*$ for any $\tilde{p}, \tilde{q} \in \mathbb{Q}$. The magnitude of $\tilde{q}$ is defined by $|\tilde{q}| = \sqrt{\tilde{q}\tilde{q}^*} = \sqrt{\tilde{q}^*\tilde{q}}$. And $\tilde{q}$ is invertible if and only if $|\tilde{q}|$ is positive. In this case, we have $\tilde{q}^{-1} = \tilde{q}^*/|\tilde{q}|$.

The set of all unit quaternions is $\mathbb{U} := \{\tilde{q} \in \mathbb{R}^4 \mid |\tilde{q}| = 1\}$, which can be regarded as a unit sphere in $\mathbb{R}^4$. Equivalently, a unit quaternion has the following form: $\tilde{q} = [\cos(\theta/2), \sin(\theta/2)\boldsymbol{n}]$, where $\boldsymbol{n} = (n_x, n_y, n_z)$ is a unit vector and θ is an angle. Let a vector $\boldsymbol{t}_1 \in \mathbb{R}^3$ rotates θ radians around axis $\boldsymbol{n}$ to reach $\boldsymbol{t}_2 \in \mathbb{R}^3$. This process can be represented by a quaternion as

$$[0, \boldsymbol{t}_2] = \tilde{q}[0, \boldsymbol{t}_1]\tilde{q}^*.$$

Using rotation matrix in $SO(3)$, we also have $\boldsymbol{t}_2 = R\boldsymbol{t}_1$, where

$$R = \cos(\theta)I_3 + (1 - \cos(\theta))\boldsymbol{n}\boldsymbol{n}^\top + \sin(\theta)\boldsymbol{n}^\wedge, \quad \text{and} \quad \boldsymbol{n}^\wedge = \begin{pmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{pmatrix}.$$

The relationship between rotation matrix and unit quaternion is given in the next lemma.

Lemma A.1. *[12] Given a unit quaternion $\tilde{q} = [q_0, \boldsymbol{q}] = [q_0, q_1, q_2, q_3] \in \mathbb{U}$ and a vector $\boldsymbol{t} \in \mathbb{R}^3$. Then $[0, R\boldsymbol{t}] = \tilde{q}[0, \boldsymbol{t}]\tilde{q}^*$, where the rotation matrix $R \in SO(3)$ satisfies*

$$R = \boldsymbol{q}\boldsymbol{q}^\top + q_0^2 I + 2q_0\boldsymbol{q}^\wedge + (\boldsymbol{q}^\wedge)^2.$$

If the rotation matrix R is compound motion of two rotations, i.e., $R = R_2R_1$, then the corresponding quaternion $\tilde{q}$ can be formulated as $\tilde{q} = \tilde{q}_2\tilde{q}_1$.

A.3. Quaternion matrix

The collections of real, complex and quaternion $m \times n$ matrices are denoted by $\mathbb{R}^{m\times n}$, $\mathbb{C}^{m\times n}$, $\mathbb{Q}^{m\times n}$, respectively.

A quaternion matrix [23] $\tilde{A} = (\tilde{a}_{ij}) \in \mathbb{Q}^{m\times n}$ can be denoted as

$$\tilde{A} = A_0 + A_1\boldsymbol{i} + A_2\boldsymbol{j} + A_3\boldsymbol{k},$$

where $A_0,\ A_1,\ A_2,\ A_3 \in \mathbb{R}^{m\times n}$. The conjugate transpose of $\tilde{A}$ is $\tilde{A}^* = (a_{ji}^*)$. If $\tilde{A} = \tilde{A}^*$, then $\tilde{A}$ is a Hermitian matrix. We denote $\mathbb{H}^{n\times n} = \{\tilde{A} \in \mathbb{Q}^{n\times n} \mid \tilde{A} = \tilde{A}^*\}$ be the set of Hermitian matrices.

For $\tilde{A},\ \tilde{B} \in \mathbb{Q}^{m\times n}$, their inner product is defined as

$$\left\langle \tilde{A}, \tilde{B} \right\rangle = \mathrm{tr}(\tilde{A}^*\tilde{B}),$$

where $\mathrm{tr}(\tilde{A}^*\tilde{B})$ denotes the trace of matrix $\tilde{A}^*\tilde{B}$. The Frobenius norm of A is given by

$$\|\tilde{A}\|_F = \sqrt{\left\langle \tilde{A}, \tilde{A} \right\rangle} = \sqrt{\mathrm{tr}(\tilde{A}^*\tilde{A})} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |\tilde{a}_{ij}|^2}.$$

In addition, we present some useful properties of quaternion matrices which generalize those in the complex case. It is important to note that, in general, $\mathrm{tr}(\tilde{A}\tilde{B}) \neq \mathrm{tr}(\tilde{B}\tilde{A})$ due to the non-commutativity of quaternion multiplication. This is a key difference from the complex matrix case, and as a result, some properties require careful consideration.

Theorem A.1. *Suppose $\tilde{A},\ \tilde{B} \in \mathbb{Q}^{n\times n}$ are two any quaternion matrix, then the following statements hold:*

(i) $(\tilde{A}\tilde{B})^* = \tilde{B}^*\tilde{A}^*$.

(ii) $\mathrm{tr}(\tilde{A}^*) = \mathrm{tr}(\tilde{A})^*$.

(iii) $\mathrm{tr}(\tilde{A}^*\tilde{B}) = \mathrm{tr}(\tilde{B}^*\tilde{A})^*$, *or* $\left\langle \tilde{A}, \tilde{B} \right\rangle = \left\langle \tilde{B}, \tilde{A} \right\rangle^*$.

(iv) If $\tilde{A} \in \mathbb{H}^{n\times n}$, then $\tilde{x}^*\tilde{A}\tilde{x}$ is a real number for any $\tilde{x} \in \mathbb{Q}^n$.

(v) $\mathrm{Re}(\mathrm{tr}(\tilde{A}\tilde{B})) = \mathrm{Re}(\mathrm{tr}(\tilde{B}\tilde{A}))$.

(vi) If $\tilde{A} \in \mathbb{H}^{n\times n}$ and $\tilde{x} \in \mathbb{Q}^n$, then $\mathrm{tr}(\tilde{x}\tilde{x}^*\tilde{A}) + \mathrm{tr}(\tilde{A}\tilde{x}\tilde{x}^*) = 2\tilde{x}^*\tilde{A}\tilde{x}$.

Proof. We omit the proofs of (i)–(v) which are not hard to verify. Here we provide a proof of (vi), a result that will be frequently employed in the subsequent analysis. From (ii) and (iii), we have

$$\mathrm{tr}(\tilde{x}\tilde{x}^*\tilde{A}) = \mathrm{tr}(\tilde{A}\tilde{x}\tilde{x}^*)^*,$$

which implies that $\mathrm{tr}(\tilde{x}\tilde{x}^*\tilde{A}) + \mathrm{tr}(\tilde{A}\tilde{x}\tilde{x}^*) = 2\,\mathrm{Re}(\mathrm{tr}(\tilde{x}\tilde{x}^*\tilde{A}))$. According to (iv) and (v), we have

$$\mathrm{Re}(\mathrm{tr}(\tilde{x}\tilde{x}^*\tilde{A})) = \mathrm{Re}(\tilde{x}^*\tilde{A}\tilde{x}) = \tilde{x}^*\tilde{A}\tilde{x},$$

which completes the proof. □

Lemma A.2. *([27] Corollary 6.2) Suppose $\tilde{X} \in \mathbb{H}^{n\times n}$, then there exists a unitary matrix $\tilde{U} \in \mathbb{Q}^{n\times n}$ (that is, $\tilde{U}\tilde{U}^* = \tilde{U}^*\tilde{U} = I$) such that*

$$\tilde{U}^*\tilde{X}\tilde{U} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n), \tag{5}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are real eigenvalues of the quaternion matrix $\tilde{X}$. Furthermore, if $\tilde{X}$ is a positive definite matrix, then $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all positive numbers.

For any Hermitian matrix $\tilde{X} \in \mathbb{H}^{n\times n}$, define its spectral norm by $\|\tilde{X}\|_2 = \max_i |\lambda_i|$. We denote the cones of quaternion Hermitian positive semidefinite and positive definite matrices by $\mathbb{H}_+^{n\times n} = \{\tilde{X} \in \mathbb{H}^{n\times n} \mid \tilde{X} \succeq 0\}$ and $\mathbb{H}_{++}^{n\times n} = \{\tilde{X} \in \mathbb{H}^{n\times n} \mid \tilde{X} \succ 0\}$, respectively.

Lemma A.3. *Let the dual cone of $\mathbb{H}_+^{n\times n}$ be*

$$(\mathbb{H}_+^{n\times n})^* = \{Y \in \mathbb{H}^{n\times n} \mid \mathrm{tr}(\tilde{X}\tilde{Y}) + \mathrm{tr}(\tilde{Y}\tilde{X}) \geq 0,\ \forall \tilde{X} \in \mathbb{H}_+^{n\times n}\}.$$

Then we have $\mathbb{H}_+^{n\times n} = (\mathbb{H}_+^{n\times n})^*$.

Proof. Given any $\tilde{Y} \in \mathbb{H}_+^{n\times n}$, by Lemma A.2, we have $\tilde{Y} = \sum_{i=1}^n \lambda_i \tilde{u}_i \tilde{u}_i^*$. Then

$$\operatorname{tr}(\tilde{X}\tilde{Y}) + \operatorname{tr}(\tilde{Y}\tilde{X}) = \sum_{i=1}^n \lambda_i \left(\operatorname{tr}(\tilde{X}\tilde{u}_i\tilde{u}_i^*) + \operatorname{tr}(\tilde{u}_i\tilde{u}_i^*\tilde{X})\right) = 2\sum_{i=1}^n \lambda_i \tilde{u}_i^*\tilde{X}\tilde{u}_i,$$

where the last equality follows from Theorem A.1. Since $\lambda_i \geq 0$, $\tilde{u}_i^*\tilde{X}\tilde{u}_i \geq 0$, we have $\tilde{Y} \in (\mathbb{H}_+^{n\times n})^*$, which implies $\mathbb{H}_+^{n\times n} \subseteq (\mathbb{H}_+^{n\times n})^*$.

Next we prove $(\mathbb{H}_+^{n\times n})^* \subseteq \mathbb{H}_+^{n\times n}$. Suppose $\tilde{Y} \notin \mathbb{H}_+^{n\times n}$. Then there exists a $\tilde{q} \in \mathbb{Q}^n$ such that $\tilde{q}^*\tilde{Y}\tilde{q} < 0$. Let $\tilde{X} = \tilde{q}\tilde{q}^* \in \mathbb{H}_+^{n\times n}$, and we have

$$\begin{aligned}\operatorname{tr}(\tilde{X}\tilde{Y}) + \operatorname{tr}(\tilde{Y}\tilde{X}) &= \operatorname{tr}(\tilde{q}\tilde{q}^*\tilde{Y}) + \operatorname{tr}(\tilde{Y}\tilde{q}\tilde{q}^*)\\ &= 2\tilde{q}^*\tilde{Y}\tilde{q} < 0.\end{aligned}$$

Hence, $\tilde{Y} \notin (\mathbb{H}_+^{n\times n})^*$ which implies $(\mathbb{H}_+^{n\times n})^* \subseteq \mathbb{H}_+^{n\times n}$. □

A.4. Distance metrics for rotations

Assume that R_1, R_2 are two rotation matrices, and $\tilde{q}_1, \tilde{q}_2$ are corresponding quaternion representations. Define the rotation angle θ_{ij} of two different rotations as

$$\theta = \operatorname{dist}_a(R_1, R_2) = \|\operatorname{Log}(R_1^\top R_2)\|_2 = \arccos\frac{\operatorname{tr}(R_1^\top R_2) - 1}{2},$$

where $\operatorname{Log}(R)$ denotes the logarithm map for $SO(3)$ and return its axis-angle representation $\boldsymbol{n}$. It can be shown that this distance is a geodesic distance [11], i.e., it is the length of the minimum path between R_1 and R_2 on the manifold $SO(3)$. Then the chordal distance and the quaternion distance between two rotations can be interpreted as

$$\operatorname{dist}_c(R_1, R_2) = \|R_1 - R_2\|_F = 2\sqrt{2}\sin(\theta/2),$$
$$\operatorname{dist}_q(R_1, R_2) = \|\tilde{q}_1 - \tilde{q}_2\|_F = 2\sin(\theta/4).$$

B. PROOFS FOR SECTION 4.1

B.1. High-dimensional probability for quaternion random matrix

Definition B.1. *A quaternion random variable $\tilde{q} \in \mathbb{Q}$ is said to follow a standard quaternion normal distribution if its components q_0, q_1, q_2, q_3 are independent and identically distributed (i.i.d.) with $q_i \sim \mathcal{N}(0, \frac{1}{4})$. Under this distribution, $\tilde{q}$ satisfies $\mathbb{E}(\tilde{q}) = 0$ and $\mathbb{E}|\tilde{q}|^2 = 1$, and we denote $\tilde{q} \sim \mathcal{QN}(0,1)$.*

Prior to the proof of Theorem 4.1, we present two lemmas as follows.

Lemma B.1. *Let $\tilde{A} \in \mathbb{H}^{n\times n}$ be a quaternion Hermitian random matrix with i.i.d. off-diagonal entries following a standard quaternion normal distribution $\mathcal{QN}(0,1)$ and zeros on the diagonal. Then we have*

$$\mathbb{P}\{\|\tilde{A}\|_2 > (2 + c_1)\sqrt{n}\} \leq e^{-c_1^2 n/2}, \tag{6}$$

where $c_1 > 0$ is an arbitrary constant.

Proof. The tail bounds for real and complex cases were established in [28, 20, 29] using Slepian's comparison theorem and Gaussian concentration. However, the non-commutativity of quaternion multiplication prevents trivial extension of these results. We divide the proof into three steps.

Step I: We first bound the largest eigenvalue of a quaternion Hermitian random matrix $\tilde{A}$. Let

$$\begin{aligned}\sigma_{\max}(\tilde{A}) &= \max_{\tilde{q}\in\mathbb{Q}^n:\|\tilde{q}\|_2=1} \tilde{q}^*\tilde{A}\tilde{q}\\ &= \max_{\tilde{q}\in\mathbb{Q}^n:\|\tilde{q}\|_2=1} \frac{1}{2}\left\langle \tilde{A}, \tilde{q}\tilde{q}^*\right\rangle + \frac{1}{2}\left\langle \tilde{q}\tilde{q}^*, \tilde{A}\right\rangle\end{aligned}$$

denote the largest eigenvalue of $\tilde{A}$. For any $\tilde{p}, \tilde{q} \in \mathbb{Q}^n$ which satisfiy $\|\tilde{p}\|_2 = \|\tilde{q}\|_2 = 1$, the real valued Gaussian process $X_{\tilde{q}} = \tilde{q}^*\tilde{A}\tilde{q}$ satisfies

$$\begin{aligned}\mathbb{E}(X_{\tilde{q}} - X_{\tilde{p}})^2 =& \frac{1}{4}\mathbb{E}\left(\left\langle \tilde{A}, \tilde{q}\tilde{q}^* - \tilde{p}\tilde{p}^*\right\rangle + \left\langle \tilde{q}\tilde{q}^* - \tilde{p}\tilde{p}^*, \tilde{A}\right\rangle\right)^2\\ =& \frac{1}{4}\mathbb{E}\left(\sum_{i,j}\tilde{a}_{ij}^*(\tilde{q}_i\tilde{q}_j^* - \tilde{p}_i\tilde{p}_j^*) + \sum_{i,j}(\tilde{q}_j\tilde{q}_i^* - \tilde{p}_j\tilde{p}_i^*)\tilde{a}_{ij}\right)^2.\end{aligned}$$

Note that, for any $1 \leq i, j \leq n$, $\tilde{A}$ is a Hermitian matrix and the last two terms are quaternion conjugates of each other. Then we have

$$\begin{aligned}\mathbb{E}(X_{\tilde{q}} - X_{\tilde{p}})^2 =& \frac{1}{4}\mathbb{E}\left(4\sum_{i<j}\mathrm{Re}(\tilde{a}_{ij}^*(\tilde{q}_i\tilde{q}_j^* - \tilde{p}_i\tilde{p}_j^*))\right)^2 \\ =& 4\sum_{i<j}\mathbb{E}\left(\mathrm{Re}(\tilde{a}_{ij}^*(\tilde{q}_i\tilde{q}_j^* - \tilde{p}_i\tilde{p}_j^*))\right)^2 \\ =& \sum_{i<j}|\tilde{q}_i\tilde{q}_j^* - \tilde{p}_i\tilde{p}_j^*|^2.\end{aligned}$$

The second equality follows from that all $a_{ij},\ i<j$ are independent and $\mathbb{E}(a_{ij}) = 0$. The last equality follows from the fact that, for any $\tilde{x} \sim \mathcal{QN}(0,1)$ and $\tilde{y} \in \mathbb{Q}$,

$$\begin{aligned}\mathbb{E}_{\tilde{x}}\left(\mathrm{Re}(\tilde{x}^*\tilde{y})\right)^2 =& \mathbb{E}_{\tilde{x}}\left(x_0y_0 + x_1y_1 + x_2y_2 + x_3y_3\right)^2 \\ =& \mathbb{E}_{\tilde{x}}\left(x_0^2y_0^2 + x_1^2y_1^2 + x_2^2y_2^2 + x_3^2y_3^2\right) \\ =& \frac{1}{4}|\tilde{y}|^2.\end{aligned}$$

Further, since for any $\tilde{x}, \tilde{y} \in \mathbb{Q}$,

$$|\tilde{x}\tilde{y}|^2 = |\tilde{x}|^2|\tilde{y}|^2, \tag{7a}$$

$$\mathrm{Re}(\tilde{x}\tilde{y}) = \mathrm{Re}(\tilde{y}\tilde{x}), \tag{7b}$$

$$\mathrm{Re}(\tilde{x}) + \mathrm{Re}(\tilde{y}) = \mathrm{Re}(\tilde{x} + \tilde{y}), \tag{7c}$$

we obtain

$$\begin{aligned}\sum_{i<j}|\tilde{q}_i\tilde{q}_j^* - \tilde{p}_i\tilde{p}_j^*|^2 =& \frac{1}{2}\sum_{i,j}|\tilde{q}_i\tilde{q}_j^* - \tilde{p}_i\tilde{p}_j^*|^2 \\ =& \frac{1}{2}\sum_{i,j}\left(|\tilde{q}_i|^2|\tilde{q}_j|^2 + |\tilde{p}_i|^2|\tilde{p}_j|^2 - 2\,\mathrm{Re}(\tilde{q}_j\tilde{q}_i^*\tilde{p}_i\tilde{p}_j^*)\right) \qquad \text{(using (7a))} \\ =& \frac{1}{2}\left(\sum_i|\tilde{q}_i|^2\sum_j|\tilde{q}_j|^2 + \sum_i|\tilde{p}_i|^2\sum_j|\tilde{p}_j|^2 - 2\sum_{i,j}\mathrm{Re}(\tilde{q}_j\tilde{q}_i^*\tilde{p}_i\tilde{p}_j^*)\right) \\ =& 1 - \mathrm{Re}(\sum_{i,j}\tilde{q}_i^*\tilde{p}_i\tilde{p}_j^*\tilde{q}_j) \qquad \text{(using (7b) and (7c))} \\ =& 1 - \mathrm{Re}\left\langle\sum_i(\tilde{p}_i^*\tilde{q}_i), \sum_j(\tilde{p}_j^*\tilde{q}_j)\right\rangle \\ =& 1 - |\tilde{p}^*\tilde{q}|^2.\end{aligned}$$

Due to $\|\tilde{p}\|_2 = \|\tilde{q}\|_2 = 1$, we have $|\tilde{p}^*\tilde{q}| \leq 1$ and

$$\begin{aligned}1 - |\tilde{p}^*\tilde{q}|^2 &\leq 2(1 - |\tilde{p}^*\tilde{q}|) \\ &\leq 2(1 - \mathrm{Re}(\tilde{p}^*\tilde{q})) \\ &= \|\tilde{p} - \tilde{q}\|_2^2.\end{aligned}$$

Therefore, we obtain $\mathbb{E}(X_{\tilde{q}} - X_{\tilde{p}})^2 \leq \|\tilde{p} - \tilde{q}\|_2^2$.

Step II: Define $Y_{\tilde{q}} = \langle g_0, p_0\rangle + \langle g_1, p_1\rangle + \langle g_2, p_2\rangle + \langle g_3, p_3\rangle$, where $g = [g_0, g_1, g_2, g_3]$ is the canonical Gaussian random vector in $\mathbb{R}^{4n}$, i.e., g_k are i.i.d. and $g_k \sim \mathcal{N}(0, I_n)$. Regarding both $\tilde{q}, \tilde{p}$ as real vectors, then we have

$$\begin{aligned}\mathbb{E}_g(Y_{\tilde{q}} - Y_{\tilde{p}})^2 &= \mathbb{E}_g\left\langle g, \tilde{q} - \tilde{p}\right\rangle^2 \\ &= \mathbb{E}_g(\tilde{q} - \tilde{p})^\top\mathbb{E}(gg^\top)(\tilde{q} - \tilde{p}) \\ &= \|\tilde{p} - \tilde{q}\|_2^2.\end{aligned}$$

Step III: According to Step I and II, we get $\mathbb{E}(X_{\tilde{q}} - X_{\tilde{p}})^2 \leq \mathbb{E}_g(Y_{\tilde{q}} - Y_{\tilde{p}})^2$. Using Slepian's comparison theorem [30], we have

$$\mathbb{E}(\sigma_{\max}(\tilde{A})) = \mathbb{E}\sup_{\|\tilde{q}\|_2=1}X_{\tilde{q}} \leq \mathbb{E}\sup_{\|\tilde{q}\|_2=1}Y_{\tilde{q}} = \mathbb{E}\|g\|_2 \leq \sqrt{\mathbb{E}\|g\|_2^2} = 2\sqrt{n}.$$

Since $\sigma_{\max}(\cdot)$ is a real valued Lipschitz function with Lipschitz constant 1, the concentration in Gauss space [31] implies that

$$\mathbb{P}\{\sigma_{\max}(\tilde{A}) - \mathbb{E}(\sigma_{\max}(\tilde{A})) \geq t\} \leq e^{-t^2/2}.$$

Let $t = c_1\sqrt{n}$, we have $\mathbb{P}\{\|\tilde{A}\|_2 > (2+c_1)\sqrt{n}\} \leq e^{-c_1^2 n/2}$. $\square$

Lemma B.2. *Let $\tilde{A} \in \mathbb{H}^{n\times n}$ be a quaternion Hermitian random matrix with i.i.d. off-diagonal entries following a standard quaternion normal distribution $\mathcal{QN}(0,1)$ and zeros on the diagonal. Then we have*

$$\mathbb{P}\{\|\tilde{A}\mathbb{1}\|_\infty > c_2\sqrt{n\log n}\} \leq (1+c_2^2)n^{3-2c_2^2}, \tag{8}$$

where $c_2 > 0$ is an arbitrary constant.

Proof. Since $(\tilde{A}\mathbb{1})_i = \sum_j \tilde{a}_{ij}$, and all $\tilde{a}_{ij}$, $i<j$ are independent, we have $(\tilde{A}\mathbb{1})_i$ also follows quaternion normal distribution, and

$$\mathbb{E}(\tilde{A}\mathbb{1})_i = 0, \quad \mathbb{E}|(\tilde{A}\mathbb{1})_i|^2 = \sum_{j=1}^n \mathbb{E}|\tilde{a}_{ij}|^2 = n-1.$$

Although $(\tilde{A}\mathbb{1})_i$ are not independent with each other, we can find an upper bound based on subadditivity in probability theory, i.e.,

$$\mathbb{P}\{\|\tilde{A}\mathbb{1}\|_\infty > t\} = \mathbb{P}\left\{\bigcup_i |(\tilde{A}\mathbb{1})_i| > t\right\} \leq \sum_i \mathbb{P}\left\{|(\tilde{A}\mathbb{1})_i| > t\right\}. \tag{9}$$

For any quaternion random variable $\tilde{q} \sim \mathcal{QN}(0,1)$, $Y = 4|\tilde{q}|^2 = (2\tilde{q}_0)^2 + (2\tilde{q}_1)^2 + (2\tilde{q}_2)^2 + (2\tilde{q}_3)^2$ follows Chi-square distribution, i.e., $Y \sim \mathcal{X}^2(4)$. The probability density function of Y is $f_Y(y) = \frac{1}{4}ye^{-y/2}$. Let $R = \sqrt{Y}$, and then random variable R follows Rayleigh distribution whose probability density function is $f_R(r) = 2rf_Y(r^2) = \frac{r^3}{2}e^{-r^2/2}$. The tail of random variable R satisfies

$$\mathbb{P}\{2|\tilde{q}| > t\} = \int_t^{+\infty} f_R(r)dr = \left(1+\frac{t^2}{2}\right)e^{-t^2/2}. \tag{10}$$

Combining (9) and (10), we have

$$\begin{aligned}\mathbb{P}\{\|\tilde{A}\mathbb{1}\|_\infty > t\} &\leq \sum_i \mathbb{P}\left\{\left|\frac{1}{\sqrt{n-1}}(\tilde{A}\mathbb{1})_i\right| > \frac{t}{\sqrt{n}}\right\} \\ &\leq \left(1+\frac{2t^2}{n}\right)ne^{-2t^2/n}.\end{aligned}$$

Let $t = c_2\sqrt{n\log n}$, and we get

$$\mathbb{P}\{\|\tilde{A}\mathbb{1}\|_\infty > c_2\sqrt{n\log n}\} \leq (1+2c_2^2\log n)n^{1-2c_2^2} \leq (1+c_2^2)n^{3-2c_2^2},$$

where the last inequality holds for any $n \geq 2$. $\square$

B.2. Proof of Theorem 4.1

Proof. Here we prove that the quaternion random matrix $\tilde{A}$ is $\tilde{q}$-discordant with high probability. Since

$$\begin{aligned}\|\operatorname{diag}(\tilde{q}^*)\tilde{A}\operatorname{diag}(\tilde{q})\|_2 &= \|\operatorname{diag}(\tilde{q})\|_2^2\|\tilde{A}\|_2 = \|\tilde{A}\|_2, \\ \|\operatorname{diag}(\tilde{q}^*)\tilde{A}\operatorname{diag}(\tilde{q})\tilde{\mathbb{1}}\|_\infty &= \|\operatorname{diag}(\tilde{q}^*)\tilde{A}\tilde{q}\|_\infty = \|\tilde{A}\tilde{q}\|_\infty,\end{aligned}$$

we have that the quaternion matrix $\tilde{A}$ is $\tilde{q}$-discordant if and only if $\operatorname{diag}(\tilde{q}^*)\tilde{A}\operatorname{diag}(\tilde{q})$ is $\tilde{\mathbb{1}}$-discordant.

Furthermore, denote $\tilde{B} = \operatorname{diag}(\tilde{q}^*)\tilde{A}\operatorname{diag}(\tilde{q})$ and we have $\tilde{b}_{ij} = \tilde{q}_i^*\tilde{a}_{ij}\tilde{q}_j$. Since $\tilde{a}_{ij}$ is a standard quaternion normal random variable which has i.i.d. real Gaussian entries, $\tilde{b}_{ij}$ also follows a normal distribution. The expectation $\mathbb{E}(\tilde{b}_{ij}) = \tilde{q}_i^*\mathbb{E}(\tilde{a}_{ij})\tilde{q}_j = 0$, and the variance $Var(\tilde{b}_{ij}) = \mathbb{E}(\|\tilde{b}_{ij}\|^2) = 1$. Therefore, $\tilde{B}$ has the same distribution as $\tilde{A}$, and we may without loss of generality assume $\tilde{q} = \tilde{\mathbb{1}}$ in the proof.

Combining Lemmas B.1 and B.2, we have $\tilde{A}$ is $\mathbb{1}$-discordant with probability at least $1 - e^{-c_1^2 n/2} - (1+c_2^2)n^{3-2c_2^2}$. The proof is completed. $\square$

B.3. Two lemmas

For any feasible point $\tilde{x} \in \mathbb{U}^n$ and $\tilde{z} \in \mathbb{U}$, since the function value $\tilde{x}^*\tilde{C}\tilde{x} = \tilde{z}^*\tilde{x}^*\tilde{C}\tilde{x}\tilde{z}$, the available data are insufficient to distinguish $\tilde{x}$ from $\tilde{x}\tilde{z}$. We denote the equivalence class of $\tilde{x}$ by

$$[\tilde{x}] = \{\tilde{x}\tilde{z} : \tilde{z} \in \mathbb{U}\}.$$

First, we prove that if the function value of $\tilde{x}$ in (Q-QP) is greater than that of the exact rotation $\tilde{q}$, then the distance between $[\tilde{x}]$ and $\tilde{q}$ under 2-norm can be bounded. In a more general setting, the bound also holds for any $\|\tilde{x}\|_2^2 = n$.

Lemma B.3. *If $\tilde{W}$ is $\tilde{q}$-discordant with constant $c_1, c_2 > 0$, and $\tilde{x} \in \mathbb{Q}^n$ satisfies $\|\tilde{x}\|_2^2 = n$ and $\tilde{q}^*\tilde{C}\tilde{q} \leq \tilde{x}^*\tilde{C}\tilde{x}$, then we have*

$$\operatorname{dist}(\tilde{q}, [\tilde{x}]) = \min_{\tilde{z}\in\mathbb{U}} \|\tilde{x}\tilde{z} - \tilde{q}\|_2 \leq 4(2 + c_1)\sigma. \tag{11}$$

Proof. First, we reformulated the distance $\operatorname{dist}^2(\tilde{q}, [\tilde{x}])$ as

$$\min_{\tilde{z}\in\mathbb{U}} \|\tilde{x}\tilde{z} - \tilde{q}\|_2^2 = 2\left(n - \min_{\tilde{z}\in\mathbb{U}} \operatorname{Re}(\tilde{z}^*\tilde{x}^*\tilde{q})\right) = 2(n - |\tilde{x}^*\tilde{q}|). \tag{12}$$

Combining $\tilde{q}^*\tilde{C}\tilde{q} \leq \tilde{x}^*\tilde{C}\tilde{x}$ with $C = \tilde{q}\tilde{q}^* + \sigma\tilde{W}$, we obtain

$$n^2 - |\tilde{x}^*\tilde{q}|^2 \leq \sigma\left(\tilde{x}^*\tilde{W}\tilde{x} - \tilde{q}^*\tilde{W}\tilde{q}\right).$$

Since $n + |\tilde{x}^*\tilde{q}| \geq n$, we have

$$n - |\tilde{x}^*\tilde{q}| \leq \sigma n^{-1}\left(\tilde{x}^*\tilde{W}\tilde{x} - \tilde{q}^*\tilde{W}\tilde{q}\right). \tag{13}$$

By using the definition of $\tilde{q}$-discordant and Theorem A.1 (vi), we have

$$\begin{aligned}
\tilde{x}^*\tilde{W}\tilde{x} - \tilde{q}^*\tilde{W}\tilde{q} =& \operatorname{Re}\left((\tilde{x} - \tilde{q})^*\tilde{W}(\tilde{x} + \tilde{q})\right) \\
\leq& \|\tilde{x} - \tilde{q}\|_2\|\tilde{W}\|_2\|\tilde{x} + \tilde{q}\|_2 \\
\leq& 2(2 + c_1)n\|\tilde{x} - \tilde{q}\|_2,
\end{aligned}$$

where the first equality follows from $\operatorname{Re}(\tilde{x}^*\tilde{W}\tilde{q}) = \operatorname{Re}(\tilde{q}^*\tilde{W}\tilde{x})$. Assume $\tilde{x}$ and $\tilde{q}$ are optimally aligned, i.e., $|\tilde{x}^*\tilde{q}| = \tilde{x}^*\tilde{q}$. Together with (12) and (13), we obtain $\|\tilde{x} - \tilde{q}\|_2^2 \leq 4(2 + c_1)\sigma\|\tilde{x} - \tilde{q}\|_2$, which implies $\|\tilde{x} - \tilde{q}\|_2 \leq 4(2 + c_1)\sigma$. ☐

In the next lemma, we establish an upper bound on $\|\tilde{x} - \tilde{q}\|_\infty$ when they are optimally aligned, i.e., $|\tilde{x}^*\tilde{q}| = \tilde{x}^*\tilde{q}$.

Lemma B.4. *If $\tilde{W}$ is $\tilde{q}$-discordant with constant $c_1, c_2 > 0$, and $\tilde{x} \in \mathbb{U}^n$ is a global optimizer of* (Q-QP) *which satisfies $|\tilde{x}^*\tilde{q}| = \tilde{x}^*\tilde{q}$, then*

$$\|\tilde{x} - \tilde{q}\|_\infty \leq 2\sigma n^{-1}\left(\|\tilde{W}\tilde{x}\|_\infty + 8(2 + c_1)^2\sigma\right). \tag{14}$$

Proof. Denote $e_i \in \mathbb{R}^n$ be the ith vector of the canonical basis, whose ith entry is 1 whereas all other entries are 0. Let $\bar{x} = \tilde{x} + (\tilde{q}_i - \tilde{x}_i)e_i$ whose ith entry is $\tilde{q}_i$ and the others are the same as $\tilde{x}$. Then $\bar{x} \in \mathbb{U}^n$ is also a feasible point of problem (Q-QP). Since $\tilde{x}$ is optimal, we have

$$\begin{aligned}
\tilde{x}^*\tilde{C}\tilde{x} &\geq \bar{x}^*\tilde{C}\bar{x} \\
&= \tilde{x}^*\tilde{C}\tilde{x} + (\tilde{q}_i - \tilde{x}_i)^*\tilde{C}_{ii}(\tilde{q}_i - \tilde{x}_i) + 2\operatorname{Re}((\tilde{q}_i - \tilde{x}_i)^*(\tilde{C}\tilde{x})_i).
\end{aligned} \tag{15}$$

Since $\tilde{C}_{ii} = 1$ and $\tilde{C} = \tilde{q}\tilde{q}^* + \sigma\tilde{W}$, we have

$$\operatorname{Re}((\tilde{q}_i - \tilde{x}_i)^*(\tilde{C}\tilde{x})_i) = \operatorname{Re}\left((\tilde{q}_i - \tilde{x}_i)^*\left(|\tilde{q}^*\tilde{x}|\tilde{q}_i + \sigma(\tilde{W}\tilde{x})_i\right)\right) \leq -\|\tilde{q}_i - \tilde{x}_i\|^2 \leq 0.$$

Then we obtain

$$\begin{aligned}
|\tilde{q}^*\tilde{x}||\tilde{x}_i - \tilde{q}_i|^2 &= 2|\tilde{q}^*\tilde{x}|(1 - \operatorname{Re}(\tilde{x}_i^*\tilde{q}_i)) \\
&= 2|\tilde{q}^*\tilde{x}|\operatorname{Re}((\tilde{q}_i - \tilde{x}_i)^*\tilde{q}_i) \\
&\leq 2\sigma\operatorname{Re}\left((\tilde{x}_i - \tilde{q}_i)^*(\tilde{W}\tilde{x})_i\right) \\
&\leq 2\sigma|\tilde{x}_i - \tilde{q}_i||(\tilde{W}\tilde{x})_i|.
\end{aligned} \tag{16}$$

This holds for all i, hence

$$|\tilde{q}^*\tilde{x}|\|\tilde{x} - \tilde{q}\|_\infty \leq 2\sigma\|\tilde{W}\tilde{x}\|_\infty.$$

According to Lemma B.3, we have $|\tilde{q}^*\tilde{x}| \geq n - 8(2+c_1)^2\sigma^2$ which implies

$$n\|\tilde{x} - \tilde{q}\|_\infty \leq 2\sigma\|\tilde{W}\tilde{x}\|_\infty + 8(2+c_1)^2\sigma^2\|\tilde{x} - \tilde{q}\|_\infty,$$

or equivalently, by using $\|\tilde{x} - \tilde{q}\|_\infty \leq 2$,

$$\|\tilde{x} - \tilde{q}\|_\infty \leq 2\sigma n^{-1}\left(\|\tilde{W}\tilde{x}\|_\infty + 8(2+c_1)^2\sigma\right).$$

□

B.4. Proof of Theorem 4.2

Proof. First, we establish an upper bound of $\|\tilde{W}\tilde{x}\|_\infty$. When $\tilde{W}$ is $\tilde{q}$-discordant with constant $c_1, c_2 > 0$, we have

$$\begin{aligned}\|\tilde{W}\tilde{x}\|_\infty &\leq \|\tilde{W}(\tilde{x} - \tilde{q})\|_\infty + \|\tilde{W}\tilde{q}\|_\infty \\ &\leq \|\tilde{W}\|_2\|\tilde{x} - \tilde{q}\|_2 + \|\tilde{W}\tilde{q}\|_\infty \\ &\leq 4(2+c_1)^2\sigma\sqrt{n} + c_2\sqrt{n\log n},\end{aligned} \tag{17}$$

where the last inequality follows from Definition 4.1. Combining Theorem 4.1, Lemma B.3 and B.4, we obtain

$$\begin{aligned}\|\tilde{x} - \tilde{q}\|_\infty &\leq 2\sigma n^{-1}\left(4(2+c_1)^2\sigma\sqrt{n} + c_2\sqrt{n\log n} + 8(2+c_1)^2\sigma\right) \\ &\leq 2\sigma n^{-1/2}\left(4(2+c_1)^2\sigma + c_2\sqrt{\log n} + 8(2+c_1)^2\sigma n^{-1/2}\right) \\ &\leq 2\sigma n^{-1/2}\left(4(\sqrt{2}+1)(2+c_1)^2\sigma + c_2\sqrt{\log n}\right)\end{aligned}$$

with probability at least $1 - e^{-c_1^2 n/2} - (1+c_2^2)n^{3-2c_2^2}$, where the last inequality holds for any $n \geq 2$. Let $c_1 = 1, c_2 = 2$, and we obtain the conclusion. □

C. PROOFS FOR SECTION 4.2

C.1. Derivation of the dual problem (Q-DSDP)

First, recall that (Q-SDP) is

$$\max_{\tilde{X}\in\mathbb{H}^{n\times n}} \operatorname{tr}(\tilde{C}\tilde{X}) + \operatorname{tr}(\tilde{X}\tilde{C}), \text{ s.t. } \operatorname{diag}(\tilde{X}) = \mathbb{1},\ \tilde{X} \succeq 0.$$

The constraint $\operatorname{diag}(\tilde{X}) = \mathbb{1}$ can be reformulated as n linear equality constraints, i.e., $\left\langle A_i, \tilde{X}\right\rangle = 1$, where $(A_i)_{ii} = 1$ and the other entries are all zeros. Since the dual cone of $\mathbb{H}_+^{n\times n}$ is itself (Lemma A.3), the Lagrangian function is

$$\begin{aligned}\mathcal{L}(\tilde{X}, \tilde{Y}, \boldsymbol{y}) &= -\operatorname{tr}(\tilde{C}\tilde{X}) - \operatorname{tr}(\tilde{X}\tilde{C}) - \operatorname{tr}(\tilde{Y}\tilde{X}) - \operatorname{tr}(\tilde{X}\tilde{Y}) + \sum_{i=1}^n y_i\left(2 - \operatorname{tr}(A_i\tilde{X}) - \operatorname{tr}(\tilde{X}A_i)\right) \\ &= 2\sum_{i=1}^n y_i - \operatorname{tr}\left((\tilde{C} + \tilde{Y} + \sum_{i=1}^n y_iA_i)\tilde{X}\right) - \operatorname{tr}\left(\tilde{X}(\tilde{C} + \tilde{Y} + \sum_{i=1}^n y_iA_i)\right).\end{aligned}$$

Then the dual problem $\max_{\boldsymbol{y}\in\mathbb{R}^n, \tilde{Y}\in\mathbb{H}_+^{n\times n}} \min_{\tilde{X}\in\mathbb{H}^{n\times n}} \mathcal{L}(\tilde{X}, \tilde{Y}, \boldsymbol{y})$ of (Q-SDP) is

$$\max_{\boldsymbol{y}\in\mathbb{R}^n, \tilde{Y}\in\mathbb{H}_+^{n\times n}} 2\boldsymbol{y}^\top\mathbb{1} \quad \text{s.t. } \tilde{Y} + \tilde{C} + \sum_{i=1}^n y_iA_i = 0.$$

Equivalently, since $\operatorname{tr}(\tilde{Y} + \tilde{C}) = \boldsymbol{y}^\top\mathbb{1}$, the dual problem can be reformulated as

$$\min_{\tilde{Y}\in\mathbb{H}^{n\times n}} \operatorname{tr}(\tilde{Y} + \tilde{C}) \quad \text{s.t. } \tilde{Y} + \tilde{C} \text{ is real diagonal and } \tilde{Y} \succeq 0. \tag{Q-DSDP}$$

C.2. Proof of Theorem 4.3

Proof. The proof follows the KKT conditions for a convex optimization problem ([32] Chapter 5). Conditions (i)–(iv) are the primal and dual feasibility, respectively. Condition (v) is complementary slackness. In addition, since the identity matrix I is strictly feasible for the primal problem, the Slater condition holds. Therefore, the KKT conditions are necessary and sufficient for global optimality.

Next, we give an explanation about Condition (v). When $\tilde{X}, \tilde{Y} \in \mathbb{H}^{n\times n}$ are the global optimizers, the strong duality holds and the optimal duality gap is zero, i.e., $\operatorname{tr}(\tilde{C}\tilde{X}) + \operatorname{tr}(\tilde{X}\tilde{C}) = 2\operatorname{tr}(\tilde{Y} + \tilde{C})$. Note that $\tilde{Y} + \tilde{C}$ is real diagonal and $\operatorname{diag}(\tilde{X}) = \mathbb{1}$, and we have

$$\begin{aligned}\operatorname{tr}(\tilde{C}\tilde{X}) + \operatorname{tr}(\tilde{X}\tilde{C}) &= \operatorname{tr}\left((\tilde{Y} + \tilde{C})\tilde{X}\right) + \operatorname{tr}\left(\tilde{X}(\tilde{Y} + \tilde{C})\right)\\ &= \operatorname{tr}(\tilde{C}\tilde{X}) + \operatorname{tr}(\tilde{X}\tilde{C}) + \operatorname{tr}(\tilde{Y}\tilde{X}) + \operatorname{tr}(\tilde{X}\tilde{Y}),\end{aligned}$$

which implies $\operatorname{tr}(\tilde{Y}\tilde{X}) + \operatorname{tr}(\tilde{X}\tilde{Y}) = 0$. From $\tilde{X}, \tilde{Y} \succeq 0$, we can decompose $\tilde{X} = \tilde{U}^*\tilde{U}$ and $\tilde{Y} = \tilde{V}^*\tilde{V}$. Then we obtain

$$\begin{aligned}0 = \operatorname{tr}(\tilde{Y}\tilde{X}) + \operatorname{tr}(\tilde{X}\tilde{Y}) &= \operatorname{tr}(\tilde{V}^*\tilde{V}\tilde{U}^*\tilde{U}) + \operatorname{tr}(\tilde{U}^*\tilde{U}\tilde{V}^*\tilde{V})\\ &= 2\operatorname{Re}(\operatorname{tr}(\tilde{V}^*\tilde{V}\tilde{U}^*\tilde{U}))\\ &= 2\operatorname{Re}(\operatorname{tr}(\tilde{V}\tilde{U}^*\tilde{U}\tilde{V}^*))\\ &= 2\|\tilde{U}\tilde{V}^*\|_F^2,\end{aligned}$$

where the third equality follows from Theorem A.1 (v). By using $\tilde{U}\tilde{V}^* = 0$, we derive $\tilde{X}\tilde{Y} = \tilde{Y}\tilde{X} = 0$ and vice versa.

The last statement follows from the results in [33]. Assume $\operatorname{rank}(\tilde{Y}) = n-1$, Theorems 9 and 10 in [33] implies that the primal solution $\tilde{X}$ is unique. Since $\tilde{X}$ is nonzero and $\tilde{X}\tilde{Y} = \tilde{Y}\tilde{X} = 0$, the rank of $\tilde{X}$ must be one. □

C.3. Riemannian submanifold and Dual certificate

In this subsection, we regard the set of unit quaternions $\mathbb{U}$ as a 3-dimensional sphere, which is a submanifold embedded in $\mathbb{Q}$. Then the quaternion quadratic programming (Q-QP) is an optimization problem with Riemannian submanifold constraints $\tilde{x} \in \mathbb{U}^n$, whose global optimizer should satisfy the first-order necessary optimality conditions over manifolds. We will then utilize these optimality conditions to identify the optimal dual variable $\tilde{Y}$ that satisfies Theorem 4.3, thereby further proving the tightness of the semidefinite relaxation.

The tangent space [34] of $\mathbb{U}$ at each $\tilde{y}$ and the projector are given respectively by

$$\begin{aligned}&\mathcal{T}_{\tilde{y}}\mathbb{U} = \{\tilde{a} \in \mathbb{Q} \mid \operatorname{Re}(\tilde{y}^*\tilde{a}) = 0\},\\ &\operatorname{Proj}_{\mathcal{T}_{\tilde{y}}\mathbb{U}} : \mathbb{Q} \to \mathcal{T}_{\tilde{y}}\mathbb{U} \quad \text{such that } \operatorname{Proj}_{\mathcal{T}_{\tilde{y}}\mathbb{U}}(\tilde{a}) = \tilde{a} - \operatorname{Re}(\tilde{y}^*\tilde{a})\tilde{y}.\end{aligned}$$

Suppose $\mathcal{M}_1, \mathcal{M}_2$ are two manifolds. Then it holds that $\mathcal{T}_{(u,v)}(\mathcal{M}_1 \times \mathcal{M}_2) = \mathcal{T}_u\mathcal{M}_1 \times \mathcal{T}_v\mathcal{M}_2$ (Proposition 3.20 in [34]). Hence the projector over the tangent space $\mathcal{T}_{\tilde{x}}\mathbb{U}^n$ at each $\tilde{x} = [\tilde{x}_1, \ldots, \tilde{x}_n]$ is given by

$$\operatorname{Proj}_{\mathcal{T}_{\tilde{x}}\mathbb{U}^n}(\tilde{v}) = \tilde{v} - \operatorname{Re}(\operatorname{ddiag}(\tilde{v}\tilde{x}^*))\tilde{x},$$

where $\operatorname{ddiag} : \mathbb{Q}^{n\times n} \to \mathbb{Q}^{n\times n}$ sets all off-diagonal entries of a matrix to zero. Reformulating the problem (Q-QP) as $\min_{\tilde{x}\in\mathbb{U}^n} f(\tilde{x}) := -\tilde{x}^*\tilde{C}\tilde{x}$, the gradient of f is given by $\nabla f(\tilde{x}) = -2\tilde{C}\tilde{x}$. Therefore, the Riemannian gradient of f is

$$\begin{aligned}\operatorname{grad} f(\tilde{x}) &= \operatorname{Proj}_{\mathcal{T}_{\tilde{x}}\mathbb{U}^n}(\nabla f(\tilde{x}))\\ &= 2\left(\operatorname{Re}(\operatorname{ddiag}(\tilde{C}\tilde{x}\tilde{x}^*)) - \tilde{C}\right)\tilde{x}.\end{aligned} \tag{18}$$

When $\tilde{x}$ constitutes a global optimizer of (Q-QP), the first-order necessary optimality condition is satisfied: $\operatorname{grad} f(\tilde{x}) = 0$. Furthermore, the second-order necessary optimality condition requires the Riemannian Hessian of f at $\tilde{x}$ to be positive semidefinite over the tangent space $\mathcal{T}_{\tilde{x}}\mathbb{U}^n$, i.e.,

$$\langle \tilde{v}, \operatorname{Hess} f(\tilde{x})[\tilde{v}]\rangle = 2\left\langle \tilde{v}, \left(\operatorname{Re}(\operatorname{ddiag}(\tilde{C}\tilde{x}\tilde{x}^*)) - \tilde{C}\right)\tilde{v}\right\rangle \geq 0, \quad \forall \tilde{v} \in \mathcal{T}_{\tilde{x}}\mathbb{U}^n. \tag{19}$$

This observation motivates the variables

$$\tilde{X} = \tilde{x}\tilde{x}^*, \quad \text{and} \quad \tilde{Y} = \operatorname{Re}(\operatorname{ddiag}(\tilde{C}\tilde{x}\tilde{x}^*)) - \tilde{C}. \tag{20}$$

It is not hard to verify that $\operatorname{diag}(\tilde{X}) = \mathbb{1}$, $\tilde{X} \succeq 0$ and $\tilde{Y} + \tilde{C}$ is real diagonal. In addition, (18) shows that $\tilde{Y}\tilde{X} = 0$. Hence, the conditions (i)–(iii) and (v) in Theorem 4.3 are all satisfied. In addition, (18) and (19) also indicate $\tilde{Y} \succeq 0$ (condition (iv)). Therefore, when $\tilde{x}$ constitutes a global optimizer of (Q-QP), $\tilde{X}$ and $\tilde{Y}$ defined in (20) are the global optimizers of (Q-SDP) and (Q-DSDP), respectively. The following lemma demonstrates that $\tilde{Y}$ must necessarily admit the form specified in (20), with no alternative representations possible for $\tilde{X} = \tilde{x}\tilde{x}^*$.

Lemma C.1. *A feasible $\tilde{X}$ with any rank is optimal for* (Q-SDP) *if and only if $\tilde{Y} = \operatorname{Re}(\operatorname{ddiag}(\tilde{C}\tilde{X})) - \tilde{C}$ is positive semidefinite.*

Proof. We first prove the necessary condition. If $\tilde{X}$ is optimal, there exists $\tilde{Y}_0 \succeq 0$ satisfying $\tilde{Y}_0 \tilde{X} = 0$ and $\tilde{Y}_0 + \tilde{C} = \tilde{D}$ where $\tilde{D}$ is diagonal. Then $(\tilde{D} - \tilde{C})\tilde{X} = \tilde{Y}_0 \tilde{X} = 0$, which implies $\tilde{D} = \mathrm{ddiag}(\tilde{C}\tilde{X})$. Thus, $\tilde{Y}_0 = \mathrm{ddiag}(\tilde{C}\tilde{X}) - \tilde{C}$.

For sufficiency, the conditions (i)–(iv) in Theorem 4.3 are immediately satisfied. Since $\tilde{x}_{ii} = 1$, it follows that

$$\mathrm{tr}(\tilde{Y}\tilde{X}) = 0, \quad \text{and} \quad \mathrm{tr}(\tilde{X}\tilde{Y}) = \mathrm{tr}\left(\mathrm{Re}(\mathrm{ddiag}(\tilde{C}\tilde{X})) - \tilde{X}\tilde{C}\right) = 0.$$

Then (v) in Theorem 4.3 holds due to $\tilde{X}, \tilde{Y} \succeq 0$. $\square$

C.4. Proof of Theorem 4.4

Proof. According to the first-order necessary optimality conditions $\mathrm{grad}\, f(\tilde{x}) = 0$, i.e, $\tilde{Y}\tilde{x} = 0$, and $\tilde{x}$ is nonzero, we have $\mathrm{rank}(\tilde{Y}) \leq n-1$. If for arbitrary $\tilde{v} \in \mathbb{Q}^n$ such that $\tilde{v} \perp \tilde{x}$ and $\tilde{v} \neq 0$, it holds that $\tilde{v}^*\tilde{Y}\tilde{v} > 0$, then the proof is completed. We divide the proof into three steps.

Step I: (The properties of $\mathrm{ddiag}(\tilde{C}\tilde{x}\tilde{x}^*)$) Since $\tilde{Y}\tilde{x} = 0$, we have $\mathrm{Re}\left((\tilde{C}\tilde{x})_i \tilde{x}_i^*\right)\tilde{x}_i = (\tilde{C}\tilde{x})_i$. Multiplying both sides by $\tilde{x}_i^*$, we conclude $(\tilde{C}\tilde{x})_i \tilde{x}_i^*$ is real. In addition, the second-order necessary optimality condition $\left\langle \tilde{v}, \tilde{Y}\tilde{v}\right\rangle \geq 0$ holds for all $\tilde{v} \in \mathcal{T}_{\tilde{x}}\mathbb{U}^n$. In particular, we choose $\tilde{v} = [0, \ldots, 0, \tilde{x}_i \boldsymbol{k}, 0, \ldots, 0] \in \mathcal{T}_{\tilde{x}}\mathbb{U}^n$ in which the i-th entry is non-zero and $\boldsymbol{k}$ is an imaginary units. Then we have

$$\left\langle \tilde{v}, \tilde{Y}\tilde{v}\right\rangle = (\tilde{C}\tilde{x})_i \tilde{x}_i^* - 1 \geq 0,$$

where the equality follows from $\tilde{Y}_{ii}$ is real and $\tilde{v}^*\tilde{v} = 1$. Thus,

$$(\tilde{C}\tilde{x})_i \tilde{x}_i^* = |(\tilde{C}\tilde{x})_i \tilde{x}_i^*| = |(\tilde{C}\tilde{x})_i| \geq 1. \tag{21}$$

Step II: (The lower bound of $\tilde{v}^*\tilde{Y}\tilde{v}$) Without loss of generality, we assume $\tilde{q}^*\tilde{x} = |\tilde{q}^*\tilde{x}|$. For any $\tilde{v} \in \mathcal{T}_{\tilde{x}}\mathbb{U}^n$, we have

$$\begin{aligned}
\tilde{v}^*\tilde{Y}\tilde{v} &= \tilde{v}^*\, \mathrm{ddiag}(\tilde{C}\tilde{x}\tilde{x}^*)\tilde{v} - \tilde{v}^*\tilde{C}\tilde{v} \\
&= \sum_{i=1}^n |\tilde{v}_i|^2 |(\tilde{C}\tilde{x})_i| - \tilde{v}^*\tilde{C}\tilde{v} \\
&= \sum_{i=1}^n |\tilde{v}_i|^2 |(\tilde{q}\tilde{q}^*\tilde{x})_i + \sigma(\tilde{W}\tilde{x})_i| - |\tilde{v}^*\tilde{q}|^2 - \sigma\tilde{v}^*\tilde{W}\tilde{v} \\
&\geq \sum_{i=1}^n |\tilde{v}_i|^2 |\tilde{q}_i(\tilde{q}^*\tilde{x}) + \sigma(\tilde{W}\tilde{x})_i| - |\tilde{v}^*(\tilde{q} - \tilde{x}) + \tilde{v}^*\tilde{x}|^2 - \sigma\|\tilde{W}\|_2\|\tilde{v}\|_2^2 \\
&\geq |\tilde{q}^*\tilde{x}|\|\tilde{v}\|_2^2 - \sigma\sum_{i=1}^n |\tilde{v}_i|^2 |(\tilde{W}\tilde{x})_i| - |\tilde{v}^*(\tilde{q} - \tilde{x})|^2 - \sigma\|\tilde{W}\|_2\|\tilde{v}\|_2^2 \\
&\geq \|\tilde{v}\|_2^2\left(|\tilde{q}^*\tilde{x}| - \|\tilde{q} - \tilde{x}\|_2^2 - \sigma\|\tilde{W}\|_2 - \sigma\frac{1}{\|\tilde{v}\|_2^2}\sum_{i=1}^n |\tilde{v}_i|^2 |(\tilde{W}\tilde{x})_i|\right) \\
&\geq \|\tilde{v}\|_2^2\left(|\tilde{q}^*\tilde{x}| - \|\tilde{q} - \tilde{x}\|_2^2 - \sigma\|\tilde{W}\|_2 - \sigma\|\tilde{W}\tilde{x}\|_\infty\right),
\end{aligned}$$

where the second equality follows from (21), and the last inequality follows from applying the weighted average. Furthermore, by (13) and Theorem 4.2, we obtain

$$|\tilde{q}^*\tilde{x}| - \|\tilde{q} - \tilde{x}\|_2^2 - \sigma\|\tilde{W}\|_2 - \sigma\|\tilde{W}\tilde{x}\|_\infty \geq n - 216\sigma^2 - 3\sigma\sqrt{n} - \sigma\|\tilde{W}\tilde{x}\|_\infty. \tag{22}$$

Step III: (The upper bound of $\|\tilde{W}\tilde{x}\|_\infty$) By (17) and let $c_1 = 1,\ c_2 = 2$, we have

$$\|\tilde{W}\tilde{x}\|_\infty \leq 36\sigma\sqrt{n} + 2\sqrt{n\log n}.$$

Then for any $\tilde{v} \in \mathcal{T}_{\tilde{x}}\mathbb{U}^n$,

$$\tilde{v}^*\tilde{Y}\tilde{v} \geq \sqrt{n}\|\tilde{v}\|_2^2 f(n), \tag{23}$$

where $f(n) := \sqrt{n} - 216\sigma^2 n^{-1/2} - 3\sigma - 36\sigma^2 - 2\sigma\sqrt{\log n}$. When $\sigma \leq \frac{1}{16}n^{1/4}$ and $n \geq 2$, we obtain $f(n) > 0$, which implies the conclusion. $\square$

7. REFERENCES

[1] Lanhui Wang and Amit Singer, "Exact and stable recovery of rotations for robust synchronization," *Information and Inference: A Journal of the IMA*, vol. 2, no. 2, pp. 145–193, 2013.

[2] Huikang Liu, Xiao Li, and Anthony Man-Cho So, "Resync: Riemannian subgradient-based robust rotation synchronization," *Advances in Neural Information Processing Systems*, vol. 36, pp. 5236–5261, 2023.

[3] Johan Fredriksson and Carl Olsson, "Simultaneous multiple rotation averaging using Lagrangian duality," in *Asian Conference on Computer Vision*. Springer, 2012, pp. 245–258.

[4] Frank Dellaert, David M Rosen, Jing Wu, Robert Mahony, and Luca Carlone, "Shonan rotation averaging: Global optimality by surfing $SO(p)^n$," in *European Conference on Computer Vision*. Springer, 2020, pp. 292–308.

[5] Reza Monir Vaghefi and R Michael Buehrer, "Cooperative joint synchronization and localization in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3615–3627, 2015.

[6] Daniel Martinec and Tomas Pajdla, "Robust rotation and translation estimation in multiview reconstruction," in *2007 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 2007, pp. 1–8.

[7] Linfei Pan, Dániel Baráth, Marc Pollefeys, and Johannes L Schönberger, "Global structure-from-motion revisited," in *European Conference on Computer Vision*. Springer, 2024, pp. 58–77.

[8] Luca Carlone, Roberto Tron, Kostas Daniilidis, and Frank Dellaert, "Initialization techniques for 3d SLAM: A survey on rotation estimation and its use in pose graph optimization," in *2015 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2015, pp. 4597–4604.

[9] David M Rosen, Luca Carlone, Afonso S Bandeira, and John J Leonard, "SE-Sync: A certifiably correct algorithm for synchronization over the special euclidean group," *The International Journal of Robotics Research*, vol. 38, no. 2-3, pp. 95–125, 2019.

[10] Yoel Shkolnisky and Amit Singer, "Viewing direction estimation in cryo-EM using synchronization," *SIAM Journal on Imaging Sciences*, vol. 5, no. 3, pp. 1088–1110, 2012.

[11] Richard I. Hartley, Jochen Trumpf, Yuchao Dai, and Hongdong Li, "Rotation averaging," *International Journal of Computer Vision*, vol. 103, pp. 267–305, 2012.

[12] Jack B Kuipers, *Quaternions and Rotation Sequences: a Primer with Applications to Orbits, Aerospace, and Virtual Reality*, Princeton University Press, 1999.

[13] VM Govindu, "Combining two-view constraints for motion estimation," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*. IEEE Comput. Soc, vol. 2.

[14] Shuyang Ling, "Solving orthogonal group synchronization via convex and low-rank optimization: Tightness and landscape analysis," *Mathematical Programming*, vol. 200, no. 1, pp. 589–628, 2023.

[15] Anders Eriksson, Carl Olsson, Fredrik Kahl, and Tat-Jun Chin, "Rotation averaging and strong duality," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 127–135.

[16] Alvaro Parra, Shin-Fang Chng, Tat-Jun Chin, Anders Eriksson, and Ian Reid, "Rotation coordinate descent for fast globally optimal rotation averaging," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 4298–4307.

[17] Huikang Liu, Man-Chung Yue, and Anthony Man-Cho So, "A unified approach to synchronization problems over subgroups of the orthogonal group," *Applied and Computational Harmonic Analysis*, vol. 66, pp. 320–372, 2023.

[18] Seyed Mehdi Nasiri, Hadi Moradi, and Reshad Hosseini, "A linear least square initialization method for 3D pose graph optimization problem," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2474–2479.

[19] Seyed-Mahdi Nasiri, Reshad Hosseini, and Hadi Moradi, "Novel parameterization for Gauss–Newton methods in 3-D pose graph optimization," *IEEE Transactions on Robotics*, vol. 37, no. 3, pp. 780–797, 2020.

[20] Afonso S Bandeira, Nicolas Boumal, and Amit Singer, "Tightness of the maximum likelihood semidefinite relaxation for angular synchronization," *Mathematical Programming*, vol. 163, pp. 145–167, 2017.

[21] Yiqiao Zhong and Nicolas Boumal, "Near-optimal bounds for phase synchronization," *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 989–1016, 2018.

[22] Mattheus Theodor Loots, *The Development of the Quaternion Normal Distribution*, University of Pretoria (South Africa), 2010.

[23] Liqun Qi, Ziyan Luo, Qing-Wen Wang, and Xinzhen Zhang, "Quaternion matrix optimization: Motivation and analysis," *Journal of Optimization Theory and Applications*, vol. 193, no. 1, pp. 621–648, 2022.

[24] Musheng Wei, Ying Li, Fengxia Zhang, and Jianli Zhao, *Quaternion matrix computations*, Nova Science Publishers, Incorporated, 2018.

[25] Reha H Tütüncü, Kim-Chuan Toh, and Michael J Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Mathematical Programming*, vol. 95, no. 2, pp. 189–217, 2003.

[26] Stephen J Sangwine and Nicolas Le Bihan, "Quaternion singular value decomposition based on bidiagonalization to a real or complex matrix using quaternion householder transformations," *Applied Mathematics and Computation*, vol. 182, no. 1, pp. 727–738, 2006.

[27] Fuzhen Zhang, "Quaternions and matrices of quaternions," *Linear algebra and its applications*, vol. 251, pp. 21–57, 1997.

[28] Afonso S Bandeira and Ramon Van Handel, "Sharp nonasymptotic bounds on the norm of random matrices with independent entries," *The Annals of Probability*, vol. 44, no. 4, pp. 2479 – 2506, 2016.

[29] Roman Vershynin, *Introduction to the non-asymptotic analysis of random matrices*, p. 210–268, Cambridge University Press, 2012.

[30] Michel Ledoux and Michel Talagrand, *Probability in Banach Spaces: isoperimetry and processes*, Springer Science & Business Media, 2013.

[31] Michel Ledoux, *The concentration of measure phenomenon*, Number 89. American Mathematical Soc., 2001.

[32] Stephen P Boyd and Lieven Vandenberghe, *Convex optimization*, Cambridge university press, 2004.

[33] Farid Alizadeh, Jean-Pierre A Haeberly, and Michael L Overton, "Complementarity and nondegeneracy in semidefinite programming," *Mathematical programming*, vol. 77, no. 1, pp. 111–128, 1997.

[34] Nicolas Boumal, *An introduction to optimization on smooth manifolds*, Cambridge University Press, 2023.