Towards Understanding The Semidefinite Relaxations of Truncated Least-Squares in Robust Rotation Search

Liangzu Peng†, Mahyar Fazlyab†, and René Vidal†

Abstract. The rotation search problem aims to find a 3D rotation that best aligns a given number of point pairs. To induce robustness against outliers for rotation search, prior work considers truncated least-squares (TLS), which is a non-convex optimization problem, and its semidefinite relaxation (SDR) as a tractable alternative. Whether this SDR is theoretically tight in the presence of noise, outliers, or both has remained largely unexplored. We derive conditions that characterize the tightness of this SDR, showing that the tightness depends on the noise level, the truncation parameters of TLS, and the outlier distribution (random or clustered). In particular, we give a short proof for the tightness in the noiseless and outlier-free case, as opposed to the lengthy analysis of prior work.

Key words. Truncated Least-Squares, Semidefinite Relaxations, Robust Rotation Search, Geometric Vision

MSC codes.

1. Introduction. Robust geometric estimation problems in computer vision have been studied for decades [29, 41]. However, the analysis of their computational complexity is not sufficiently well understood [52]: There are fast algorithms that run in real time [22, 44, 58, 47], and there are computational complexity theorems that negate the existence of efficient algorithms [52, 4]. For example, the commonly used consensus maximization formulation (for robust fitting) is shown to be NP hard in general [52], and its closely related truncated least-squares formulation is not approximable [4], even though they are both highly robust to noise and outliers. Between these “optimistic” algorithms and “pessimistic” theorems, semidefinite relaxations of truncated least-squares [36, 59, 60] strike a favorable balance between efficiency (as they are typically solvable in polynomial time) and robustness (which is inherited to some extent from the original formulation).

Even though noise and outliers are ubiquitous in geometric vision, and non-convex formulations and their semidefinite relaxations have been widely used in a large body of papers [34, 2, 25, 19, 42, 16, 15, 35, 8, 9, 10, 48, 28, 1, 38, 63, 26, 27, 3, 50], much fewer works [17, 43, 24, 49, 31, 57, 62, 53, 39] provide theoretical insights on the robustness of semidefinite relaxations to noise, only a few semidefinite relaxations [14, 36, 59, 60] are empirically robust to outliers, and only one paper [56] on rotation synchronization gives theoretical guarantees for noise, outliers, and both. Complementary to the story of [52] and inheriting the spirit of [56], in this paper we question whether “a specific semidefinite relaxation” for “robust rotation search” is “tight” or not, and provide tightness characterizations that account for the presence of noise, outliers, and both.

More formally, in this paper we consider the following problem (see [45, 44, 59] for what

---

*This work is an extension of our conference paper [46].

Funding: This work was supported by grants NSF 1704458, NSF 1934979 and ONR MURI 503405-78051.

†Mathematical Institute for Data Science, Johns Hopkins University (lpeng25,mahyarfazlyab,vidal@jhu.edu).

The catch is that the fast methods might not always be correct (e.g., at extreme outlier rates).

[13, 5, 64, 40] analyzed SDRs under noise but they are not for geometric vision problems.
Problem 1 (Robust Rotation Search). Let \( \{(y_i, x_i)\}_{i=1}^{\ell} \) be a collection of \( \ell \) 3D point pairs. Assume that a subset \( \mathcal{I}^* \subseteq \{1, \ldots, \ell\} \) of these pairs are related by a 3D rotation \( R_0^* \in SO(3) \) up to bounded noise \( \{e_i : \|e_i\|_2 \leq \delta\}_{i=1}^{\ell} \subseteq \mathbb{R}^3 \) with \( \delta \geq 0 \), i.e.,

\[
\begin{align*}
  y_i &= R_0^* x_i + e_i, & i &\in \mathcal{I}^* \\
  y_i &\text{ and } x_i \text{ are arbitrary} & i &\notin \mathcal{I}^*.
\end{align*}
\]

(1.1)

Here, \( \mathcal{I}^* \) is called the inlier index set. If \( i \in \mathcal{I}^* \) then \( x_i, y_i \), or \( (y_i, x_i) \) is called an inlier, otherwise it is called an outlier. The goal is to find \( R_0^* \) and \( \mathcal{I}^* \) from \( \{(y_i, x_i)\}_{i=1}^{\ell} \).

To solve this problem, we consider the truncated least-squares formulation (rotation version), where the hyper-parameter \( c_i^2 \geq 0 \) is called the truncation parameter:

(TLS-R) \( \min_{R_0 \in SO(3)} \sum_{i=1}^{\ell} \min \left\{ \|y_i - R_0 x_i\|_2^2, c_i^2 \right\} \)

While (TLS-R) is highly robust to outliers and noise [60], it is non-convex and hard to solve. Via a remarkable sequence of algebraic manipulations, [59] showed that (TLS-R) is equivalent to some non-convex quadratically constrained quadratic program (QCQP), which can be relaxed as a semidefinite program (SDR) via the standard lifting technique. The exact forms of (QCQP) and (SDR) will be shown in section 2. One approach to study how much robustness (SDR) inherits from (TLS-R) or (QCQP) is to verify if the solution of (SDR) leads to a global minimizer to (QCQP). Informally, if this is true, then we say that (SDR) is tight (cf. Definition (2.3)). Here, we make the following contributions:

- For noiseless point sets without outliers (\( e_i = 0, \mathcal{I}^* = \{1, \ldots, \ell\} \) in Problem 1), we prove that (SDR) is always tight (Theorem 3.1). While this result had already been proven in [59, Section E.3], our proof is simpler and shorter.
- For noiseless point sets with outliers, Theorem 3.2 states that (SDR) is tight for sufficiently small truncation parameters \( c_i^2 \) and random outliers (regardless of the number of outliers), but it is not tight if \( c_i^2 \) is set too large. Theorem 3.4 reveals that (SDR) is vulnerable to (e.g., not tight in the presence of) clustered outlier point pairs that are defined by a rotation different from \( R_0^* \). Different from Theorem 3.1, outliers and improper choices of \( c_i^2 \) might actually undermine the tightness of (SDR).
- For noisy point sets without outliers, Theorem 3.7, with a technical proof, shows that (SDR) is tight for sufficiently small noise and for sufficiently large \( c_i^2 \).
- The case of noisy data with outliers is the most challenging, but from our analysis of the two previous cases, a tightness characterization for this difficult case follows (Theorem 3.9). Thus, we will discuss this case only sparingly.

Paper Organization. In section 2 we review the derivations of (SDR) from (TLS-R) [59], while we also provide new insights. In section 3, we discuss our main results. In section 5, we present limitations of our work and potential avenues for future research.

Notations and Basics. We employ the MATLAB notation \([a_1; \ldots; a_\ell]\) to denote concatenation into a column vector. Given a \( 4(\ell+1) \times 4(\ell+1) \) matrix \( \mathcal{A} \), we employ the bracket notation
Subsection 2.1. Derivation. While subsection 2.1 follows the development of \cite{59} in spirit, our derivation is simpler. For example, we dispensed with the use of quaternion product \cite{30} in \cite{59}, which is a sophisticated algebraic operation. That said, it is safe to treat our (SDR) as equivalent to the naive relaxation of \cite{59}; see the appendix for a detailed discussion.

In subsection 2.2 we discuss the KKT optimality conditions that are essential for studying the interplay among (QCQP), (SDR), and (D) and the tightness of (SDR).

2.1. Derivation. The term $\|y_i - R_0x_i\|_2^2 = \|y_i\|_2^2 + \|x_i\|_2^2 - 2y_i^\top R_0 x_i$ of (TLS-R) depends linearly on the rotation $R_0$. Moreover, each entry of any rotation $R_0$ depends quadratically on its unit quaternion representation $w_0 \in S^3$; recall (1.2). One then naturally asks whether $\|y_i - R_0x_i\|_2^2$ is a quadratic form in $w_0$; the answer is affirmative:
Lemma 2.1 (Rotations and Unit Quaternions). Let $R_0$ be a 3D rotation, then we have

\[(2.1) \quad \|y_i - R_0x_i\|_2^2 = w_0^T Q_i w_0,\]

where $Q_i$ is a $4 \times 4$ positive semidefinite matrix, and $w_0 \in S^3$ is the unit quaternion representation of $R_0$. Moreover, the eigenvalues of $Q_i$ are respectively

\[(2.2) \quad (\|y_i\|_2 + \|x_i\|_2)^2, (\|y_i\|_2 + \|x_i\|_2)^2, (\|y_i\|_2 - \|x_i\|_2)^2, (\|y_i\|_2 - \|x_i\|_2)^2.\]

Proof of Lemma 2.1. To simplify notations, here, and only here, we drop all indices and consider any 3D rotation $R$ and any 3D point pairs $(y, x)$. We will first show

\[(2.3) \quad \|y - Rx\|_2^2 = w^T Q(y, x) w,\]

where $w$ is the unit quaternion representation of $R$, and $Q : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^{4 \times 4}$ is some function that maps $(y, x)$ to a $4 \times 4$ matrix $Q(y, x)$. Recall (1.2), and we note that the first, second, and third entries of $Rx$ can be written as $w' X_1 w$, $w' X_2 w$, and $w' X_3 w$, respectively, where $X_1$, $X_2$, $X_3$ are symmetric matrices respectively defined as:

\[
X_1 = \begin{bmatrix} x_1 & 0 & x_3 & -x_2 \\ 0 & x_1 & x_2 & x_3 \\ x_3 & x_2 & 0 & -x_1 \\ -x_2 & x_3 & 0 & x_1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} x_2 & -x_3 & 0 & x_1 \\ -x_3 & x_2 & x_1 & 0 \\ 0 & x_1 & x_2 & x_3 \\ x_1 & 0 & x_3 & -x_2 \end{bmatrix}, \quad X_3 = \begin{bmatrix} x_3 & x_2 & -x_1 & 0 \\ x_2 & -x_3 & 0 & x_1 \\ -x_1 & 0 & -x_3 & x_2 \\ 0 & x_1 & x_2 & x_3 \end{bmatrix}
\]

Define $U(y, x) := y_1 X_1 + y_2 X_2 + y_3 X_3$, then $y^T R x = w^T U(y, x) w$. With

\[(2.4) \quad Q(y, x) := (\|y\|_2^2 + \|x\|_2^2) I_4 - 2U(y, x),\]

we have proved (2.3). Note that, as the above derivation implies, given any unit quaternion $w' \in S^3$, there is a unique 3D rotation $R'$ which satisfies $(w')^T Q(y, x) w' = \|y - R' x\|_2^2$; the latter term $\|y - R' x\|_2^2$ is always nonnegative. Combining this with the fact that $Q(y, x)$ is symmetric, we get that $Q(y, x)$ is positive semidefinite.

It remains to find the eigenvalues of $Q(y, x)$. Note that to minimize $w^T Q(y, x) w$ over $w \in S^3$ is to minimize $\|y - R x\|_2^2$ over $R \in SO(3)$. Combining this with a geometric fact gives

\[(2.5) \quad \min_{w \in S^3} w^T Q(y, x) w = \min_{R \in SO(3)} \|y - R x\|_2^2 = (\|y\|_2^2 - \|x\|_2^2)^2,
\]

So $Q(y, x)$ has $\|y\|_2^2 - \|x\|_2^2$ as its minimum eigenvalue. Since the minimum $(\|y\|_2^2 - \|x\|_2^2)^2$ of (2.5) is attained at (at least) two different rotations $R'$'s which make $Rx$ and $y$ parallel and pointing to the same direction, i.e., there are (at least) two different unit quaternions $\pm w$ and $\pm w'$ corresponding to the minimum eigenvalue $\|y\|_2^2 - \|x\|_2^2$. Thus the eigenspace of $Q(y, x)$ associated with eigenvalue $\|y\|_2^2 - \|x\|_2^2$ is of dimension at least 2. Repeat the above argument for the maximum eigenvalue $\|y\|_2^2 + \|x\|_2^2$, and note that $Q(y, x)$ is of size $4 \times 4$, then we know all eigenvalues of $Q(y, x)$. With $Q_i := Q(y_i, x_i)$, we finish the proof.
While the exact form of $Q_i$ is complicated, Lemma 2.1 informs us of a characterization on eigenvalues of $Q_i$, which is much easier to work with. Note that, while the relation between 3D rotations and unit quaternions is well known (see, e.g., [30]), we have not found (2.2) in the literature, except in the appendix of our prior work [47].

From Lemma 2.1, we now see that (TLS-R) is equivalent to

$$(\text{TLS-Q}) \quad \min_{w_0 \in S^3} \sum_{i=1}^\ell \min \left\{ w_0^\top Q_i w_0, \ c_i^2 \right\}.$$ 

Using the following simple equality ($\theta \in \{-1, 1\}$ in [36, 59]; see also [32])

$$\min\{a, b\} = \min_{\theta \in \{0, 1\}} \theta a + (1 - \theta)b = \min \theta a + (1 - \theta)b,$$

problem (TLS-Q) can be equivalently written as

$$(2.6) \quad \min_{w_0 \in S^3, \theta_i^2 = \theta_i} \sum_{i=1}^\ell \left( \theta_i w_0^\top Q_i w_0 - \theta_i c_i^2 \right) + \sum_{i=1}^\ell c_i^2$$

Note that, while the constant $\sum_{i=1}^\ell c_i^2$ in (2.6) can be ignored, keeping it there will simplify matters. Even though the objective of (2.6) is a cubic polynomial in entries of the unknowns $w_0$ and $\theta_i$’s, problem (2.6) is equivalent to a quadratic program. Indeed, let $w_i := \theta_i w_0$, which implies $\theta_i = w_0^\top w_i$. Then (2.6) becomes

$$(2.7) \quad \min_{w_0 \in S^3, \ w_i \in \{w_0, 0\}} \sum_{i=1}^\ell \left( w_0^\top (Q_i - c_i^2 I_4) w_i \right) + \sum_{i=1}^\ell c_i^2$$

Problem (2.7) now has its objective function quadratic, while in fact its constraints are also quadratic. To see this, one easily verifies that the binary constraint $w_i \in \{w_0, 0\}$ can be equivalently written quadratically as $w_i w_0^\top = w_i w_i^\top$. Collecting all vectors of variables into a $4(\ell + 1)$ dimensional column vector $w = [w_0; \ldots; w_\ell]$, we have

$$(2.8) \quad \begin{cases} w_0 \in S^3 \\ w_i \in \{w_0, 0\} \end{cases} \iff \begin{cases} \text{tr}(w_0 w_0^\top) = 1 \\ w_i w_0^\top = w_i w_i^\top \end{cases} \iff \begin{cases} \text{tr}([ww^\top]_{00}) = 1 \\ [ww^\top]_{ii} = [ww^\top]_{ii} \end{cases}$$

In the last equivalence of (2.8) we used the notation $[\cdot]_{ij}$ of section 1. Having confirmed (2.8), we can now equivalently transform (2.7) into the following (QCQP):

$$(\text{QCQP}) \quad \begin{align*} \min_{w \in \mathbb{R}^{4(\ell + 1)}} & \quad \text{tr} \left( Q w w^\top \right) + \sum_{i=1}^\ell c_i^2 \\
\text{s.t.} & \quad [ww^\top]_{0i} = [ww^\top]_{ii}, \quad \forall \ i \in \{1, \ldots, \ell\} \\
& \quad \text{tr} \left( [ww^\top]_{00} \right) = 1 \end{align*}$$
In (QCQP), \( Q \) is our \( 4(\ell + 1) \times 4(\ell + 1) \) data matrix, symmetric and satisfying
\[
\begin{cases}
[Q]_{0i} = [Q]_{i0} = \frac{1}{2}(Q_i - c_i^2 I_4), & \forall i \in \{1, \ldots, \ell\} \\
\text{all other entries of } Q \text{ are zero.}
\end{cases}
\]

It is now not hard to derive the semidefinite relaxation (SDR) and dual program (D) from (QCQP) via lifting and standard Lagrangian calculation respectively:

**Lemma 2.2 ((SDR) and (D)).** The dual and semidefinite relaxation of (QCQP) are

\[
\begin{align*}
\text{(D)} & \quad \max_{\mu, D} \mu + \sum_{i=1}^\ell c_i^2 \quad \text{s.t.} \quad Q - \mu B - D \succeq 0 \\
\text{(SDR)} & \quad \min_{W \succeq 0} \text{tr}(QW) + \sum_{i=1}^\ell c_i^2 \\
& \quad \text{s.t.} \quad [W]_{0i} = [W]_{ii}, \quad \forall i \in \{1, \ldots, \ell\}, \\
& \quad \text{tr}([W]_{00}) = 1
\end{align*}
\]

In (D), \( B \in \mathbb{R}^{4(\ell+1) \times 4(\ell+1)} \) is a matrix of zeros except \([B]_{00} = I_4\), while \( D \in \mathbb{R}^{4(\ell+1) \times 4(\ell+1)} \) is a matrix of dual variables accounting for the \( \ell \) constraints \([ww^\top]_{00} = [ww^\top]_{ii}, \) i.e., \( D \) satisfies
\[
\begin{cases}
\mathcal{D} \text{ is symmetric, } & [D]_{ii} + 2[D]_{0i} = 0, \quad \forall i \in \{1, \ldots, \ell\} \\
\text{all other entries of } D \text{ are zero.}
\end{cases}
\]

**Proof of Lemma 2.2.** Lemma 2.2 is fairly standard, and we provide a concise proof for completeness. The Lagrangian of (QCQP) is given as
\[
\mathcal{L}_{\text{QCQP}}(w, \mu, D) = \text{tr}\left((Q - \mu B - D)ww^\top\right) + \mu + \sum_{i=1}^\ell c_i^2,
\]
where \( \mu \in \mathbb{R} \) is the Lagrangian multiplier that accounts for constraint \( \text{tr}([ww^\top]_{00}) = 1 \), and \( B \) and \( D \in \mathbb{R}^{4(\ell+1) \times 4(\ell+1)} \) have the said form. The dual function \( g_D(\mu, D) \) then is
\[
g_D(\mu, D) = \inf_W \mathcal{L}_{\text{QCQP}}(w, \mu, D) = \begin{cases} 
\mu + \sum_{i=1}^\ell c_i^2 & Q - \mu B - D \succeq 0 \\
-\infty & \text{otherwise}
\end{cases}
\]

The dual problem of (QCQP) is thus (D), where we maximize \( g_D \) over the dual variables \( \mu \) and \( D \). Next, we show that (SDR) is the dual of (D). The Lagrangian of (D) is
\[
\mathcal{L}_D(W, \mu, D) = \text{tr}\left((Q - \mu B - D)W\right) + \mu + \sum_{i=1}^\ell c_i^2.
\]
For weak duality to hold, that is for \( \mathcal{L}_D(W, \mu, D) \) to majorize \( g_D(\mu, D) \) at any feasible dual points \((\mu, D)\) with \( Q - \mu B - D \succeq 0 \), we need to add an extra constraint \( W \succeq 0 \). Then, maximizing \( \mathcal{L}_D(W, \mu, D) \) for any \( \mu \in \mathbb{R} \) and any \( D \) of the form (2.10) and then minimizing the resulting dual function of (D) over \( W \succeq 0 \) gives (SDR) as the dual program of (D).
2.2. The Tightness and Optimality Conditions. We start with the following definition:

Definition 2.3 (Tightness). (SDR) is said to be tight if it admits \( \hat{w}(\hat{w})^{\top} \) as a global minimizer, where \( \hat{w} \in \mathbb{R}^{k+1} \) globally minimizes (QCQP).

We now proceed with some basic observations. First, weak duality between (QCQP) and (D) holds as a result of Lagrangian calculation. Second:

Proposition 2.4. Strong duality between (SDR) and (D) holds.

Proof. It suffices to exhibit a strictly feasible point \((\mu, \mathcal{D})\) of (D) (such that \(Q - \mu \mathcal{B} - \mathcal{D} > 0\)) and then invoke Slater’s condition [7]. Let \(\mathcal{D}\) be of the form (2.10) satisfying \([\mathcal{D}]_{0i} := \frac{1}{2}(Q_i + c_i^2 I_4)\). Then \(Q - \mu \mathcal{B} - \mathcal{D} > 0\) is equivalent to

\[
[z_0; \ldots; z_{\ell}]^{\top} (Q - \mu \mathcal{B} - \mathcal{D}) [z_0; \ldots; z_{\ell}] > 0, \quad \forall z_i \in \mathbb{R}^4
\]

\[
-\mu \|z_{0}\|^2_2 - \sum_{i=1}^{\ell} c_i^2 \|z_{i}\|^2_2 + \sum_{i=1}^{\ell} (c_i^2 \|z_{0} - z_{i}\|^2_2 + z_{i}^{\top} Q_i z_{i}) > 0, \quad \forall z_i \in \mathbb{R}^4
\]

Since Lemma 2.1 implies \(Q_i \succeq 0\), the above holds for sufficiently small \(\mu\). Such \((\mu, \mathcal{D})\) is strictly feasible, thus finishing the proof.

Since the strong duality between (SDR) and (D) holds and every feasible point of (2.7) is also feasible for (SDR), we can then obtain the following necessary and sufficient optimality condition for characterizing the tightness of (SDR):

Proposition 2.5 (Necessary and Sufficient Optimality Conditions). Suppose that (TLS-Q) preserves all inliers \(Q_1, \ldots, Q_{k^*}\) and rejects all outliers \(Q_{k^*+1}, \ldots, Q_{\ell}\) outliers. Recall that \(\hat{w}_0\) denotes a global minimizer of (TLS-Q). (SDR) is tight if and only if there is a matrix \(\mathcal{D}\) of the form (2.10) that satisfies the following conditions:

(O1) Stationarity Condition: \[
\begin{cases}
(2[\mathcal{D}]_{0i} + Q_i - c_i^2 I_4) \hat{w}_0 = 0, & \forall i \in \{1, \ldots, k^*\} \\
(2[\mathcal{D}]_{0j} + c_j^2 I_4 - Q_j) \hat{w}_0 = 0, & \forall j \in \{k^* + 1, \ldots, \ell\}
\end{cases}
\]

(O2) Dual Feasibility: \(-\hat{\mu} \|z_{0}\|^2_2 + \sum_{i=1}^{\ell} z_{i}^{\top} [\mathcal{D}]_{0i} z_{i} - \sum_{i=1}^{\ell} z_{i}^{\top} (2[\mathcal{D}]_{0i} - Q_i + c_i^2 I_4) z_{i} \geq 0, \forall z_i \in \mathbb{R}^4
\]

(O3) Objective Value Exactness: \[
\hat{\mu} = \sum_{i=1}^{k^*} (\hat{w}_0^{\top} Q_i \hat{w}_0 - c_i^2)
\]

Proof of Proposition 2.5. Since (TLS-Q) preserves all inliers \(Q_1, \ldots, Q_{k^*}\) and rejects all outliers, \(\hat{w} = [\hat{w}_0; \ldots; \hat{w}_0; 0; \ldots; 0]\), where \(\hat{w}_0\) appeared \(k^* + 1\) times, is a global minimizer of (QCQP) and reaches the minimum value \(\sum_{i=1}^{k^*} \hat{w}_0^{\top} Q_i \hat{w}_0 + \sum_{j=k^*+1}^{\ell} c_j^2\). Clearly, \(\hat{w}^{\top} \hat{w}\) is a feasible point of (SDR). Since strong duality between (SDR) and (D) holds (Proposition 2.4), the minimum of (SDR) is \(\hat{\mu} + \sum_{i=1}^{\ell} c_i^2\), and (SDR) is tight (i.e., \(\hat{w}^{\top} \hat{w}\) globally minimizes (SDR)) if and only if the minima of (SDR) and (QCQP) coincide) the KKT conditions

\[
(Q - \hat{\mu} \mathcal{B} - \mathcal{D}) \hat{w}^{\top} = 0, \quad Q - \hat{\mu} \mathcal{B} - \mathcal{D} \succeq 0
\]
hold true. Via basic algebra and calculation, we obtain that (SDR) is tight if and only if

\[(Q - \hat{\mu} B - \hat{\mathcal{D}}) \hat{w} = 0, \quad Q - \hat{\mu} B - \hat{\mathcal{D}} \succeq 0, \quad \hat{\mu} = \sum_{i=1}^{k^*} (w_0^\top Q_i w_0 - c_i^2).\]

With \(\hat{\mu} = \sum_{i=1}^{k^*} (w_0^\top Q_i w_0 - c_i^2)\), the first condition \((Q - \hat{\mu} B - \hat{\mathcal{D}}) \hat{w} = 0\) is equivalent to

\[
\begin{align*}
\sum_{i=1}^{k^*} \frac{Q_i - c_i^2 I}{2} \hat{w}_0 &- \sum_{i=1}^{k^*} (w_0^\top Q_i w_0 - c_i^2) \hat{w}_0 - \sum_{i=1}^{k^*} [\mathcal{D}]_{0i} \hat{w}_0 = 0 \\
\frac{Q_i - c_i^2 I}{2} \hat{w}_0 - [\mathcal{D}]_{0i} \hat{w}_0 &- [\mathcal{D}]_{ki} \hat{w}_0 = 0, \quad \forall i \in \{1, \ldots, k^*\} \\
\frac{Q_i - c_i^2 I}{2} \hat{w}_0 - [\mathcal{D}]_{kj} \hat{w}_0 & = 0, \quad \forall j \in \{k^* + 1, \ldots, \ell\}
\end{align*}
\]

(2.11)

Note that our assumption on (TLS-Q) implies that \(\hat{w}_0\) is an eigenvector of \(\sum_{i=1}^{k^*} Q_i\) corresponding to its minimum eigenvalue, i.e., \(\sum_{i=1}^{k^*} Q_i \hat{w}_0 = (\sum_{i=1}^{\ell} \hat{w}_0^\top Q_i \hat{w}_0) \hat{w}_0\). Also note that \(\mathcal{D}\) is symmetric with \([\mathcal{D}]_{ii} + 2[\mathcal{D}]_{0i} = 0\) (2.10). The above (2.11) can then be simplified into

\[
\begin{align*}
\sum_{i=1}^{k^*} \frac{Q_i - c_i^2 I}{2} \hat{w}_0 &+ \sum_{i=1}^{k^*} [\mathcal{D}]_{0i} \hat{w}_0 = 0 \\
\frac{Q_i - c_i^2 I}{2} \hat{w}_0 &+ [\mathcal{D}]_{0i} \hat{w}_0 = 0, \quad \forall i \in \{1, \ldots, k^*\} \\
\frac{Q_i - c_i^2 I}{2} \hat{w}_0 &- [\mathcal{D}]_{kj} \hat{w}_0 = 0, \quad \forall j \in \{k^* + 1, \ldots, \ell\}
\end{align*}
\]

which is equivalent to (O1). The second condition \(Q - \hat{\mu} B - \hat{\mathcal{D}} \succeq 0\) is equivalent to

\[
[z_0; \ldots; z_\ell]^\top (Q - \hat{\mu} B - \hat{\mathcal{D}}) [z_0; \ldots; z_\ell] \succeq 0, \quad \forall z_i \in \mathbb{R}^4,
\]

which is the same as (O2) by rewriting using the definitions of \(Q\) (2.9) and \(\mathcal{D}\) (2.10).

Armed with Proposition 2.5, to show (SDR) is tight or not we need to find the dual certificates \([\hat{\mathcal{D}}]_{0i}\)'s (and \(\hat{\mu}\)) that fulfill the (simplified) optimality conditions. Identifying eligible \([\hat{\mathcal{D}}]_{0i}\)'s or showing that such \([\hat{\mathcal{D}}]_{0i}\)'s do not exist is a core idea in proving our main results, which are to be discussed in greater detail in the next section.

3. Tightness Characterizations. We now present our results regarding the tightness of (SDR). Subsection 3.1 treats the simplest noiseless + outlier-free case (Theorem 3.1). Subsections 3.2 and 3.3 take outliers (Theorems 3.2 and 3.4) and noise (Theorem 3.7) into account respectively; subsection 3.4 brings them together for the noisy + outliers case (Theorem 3.9).

3.1. The Noiseless + Outlier-Free Case.

Theorem 3.1 (Noiseless and Outlier-Free Point Sets). (SDR) is tight in the absence of noise and outliers, meaning that \(w^*(w^*)^\top\) globally minimizes (SDR), where \(w^* = [w_0^*; \ldots; w_\ell^*] \in \mathbb{R}^{4(\ell+1)}\) is a global minimizer of (QCQP).

Proof. Note that \(w^* = [w_0^*; \ldots; w_\ell^*]\) is a global minimizer of (QCQP) that results in the optimal value 0. Let \(\mathcal{D}\) satisfy the constraint (2.10) with \([\mathcal{D}]_{0i} := \frac{1}{2}(Q_i + c_i^2 I_4)\) for every \(i = 1, \ldots, \ell\) and let \(\hat{\mu} := -\sum_{i=1}^\ell c_i^2\). Then, with \(Q_i w_0^* = 0\) (Lemma 2.1), one easily verifies that
optimality conditions (O1) and (O3) of Proposition 2.5 hold. It remains to prove condition (O2). Substitute the values of \( [\mathcal{D}]_{\mu} \), \( \hat{\mu} \) into (O2) and it simplifies:

\[
\sum_{i=1}^{\ell} c_i^2 \| z_0 \|_2^2 + \sum_{i=1}^{\ell} z_i^T (Q_i + c_i^2 I_4) z_i - 2 \sum_{i=1}^{\ell} c_i^2 z_i^T z_0 \geq 0, \quad \forall z_i \in \mathbb{R}^4
\]

(3.1)

\[
\iff \sum_{i=1}^{\ell} \left( c_i^2 \| z_0 - z_i \|_2^2 + z_i^T Q_i z_i \right) \geq 0, \quad \forall z_i \in \mathbb{R}^4
\]

Thus (O2) holds, as every \( Q_i \) is positive semidefinite (Lemma 2.1). One also observes that the equality is attained if and only if \( z_0 = \cdots = z_\ell = w_0^* \) or \( z_0 = \cdots = z_\ell = 0 \).

Our contribution here is a shorter proof for Theorem 3.1 than [59]. Besides Lemma 2.1, another key idea that shortens the proof is our construction of the dual certificate \( \tilde{\mathcal{D}} \) (or \( [\mathcal{D}]_{\mu} \)'s). While constructing dual certificates might be an art as there might not exist general approaches for doing so, our experience is to (1) start with the simplest case (e.g., noiseless + outlier-free), (2) make observations: observe the optimality conditions (cf. Proposition 2.5), inspect the first and second order Riemannian optimality conditions (cf. [6]), discover some properties of data (e.g., Lemma 2.1), (3) repeatedly try different choices of certificates.

### 3.2. The Noiseless + Outliers Case

Different from Theorem 3.1, the tightness of (SDR) in the presence of outliers depends on the truncation parameters \( c_i^2 \):

**Theorem 3.2 (Noiseless Point Sets with Outliers).** Suppose there is no noise. Consider (TLS-Q) with outliers \( Q_{k^*+1}, \ldots, Q_{\ell} \) \( (k^* < \ell) \). Recall \( w_0^* \) denotes the unit quaternion that represents the ground-truth rotation \( R_0^* \). Let \( w^* := [w_0^*; \ldots, w_0^*; 0; \ldots; 0] \in \mathbb{R}^{4(\ell+1)} \), where \( w_0^* \) appeared \( k^* + 1 \) times, and let \( W^* := w^* (w^*)^\top \). Then we have:

- If \( 0 < c_j^2 < \lambda_{\min}(Q_j), \forall j = k^* + 1, \ldots, \ell \), then (SDR) is always tight, admitting \( W^* \) as a global minimizer.

- Suppose \( c_j^2 > (w_0^*)^\top Q_j w_0^* \) for some \( j \in \{k^* + 1, \ldots, \ell\} \). Then \( W^* \) is not a global minimizer of (SDR).

**Proof of Theorem 3.2.** For the first part, note that \( c_j^2 < \lambda_{\min}(Q_j), \forall j = k^* + 1, \ldots, \ell \), so (TLS-Q) rejects all outliers and preserves all inliers, \( w_0^* \) globally minimizes (TLS-Q), and \( w^* \) globally minimizes (QCQP) with the minimum value \( \sum_{j=k^*+1}^{\ell} c_j^2 \). Let \( [\mathcal{D}]_{\mu} := \frac{1}{2} (Q_i + c_i^2 I_4) \) \( (\forall i = 1, \ldots, k^*) \), \( [\mathcal{D}]_{\mu} := \frac{1}{2} (Q_j - c_j^2 I_4) \) \( (\forall j = k^* + 1, \ldots, \ell) \), and \( \hat{\mu} := -\sum_{i=1}^{k^*} c_i^2 \). With \( Q_i w_0^* = 0, \forall i = 1, \ldots, k^* \), (Lemma 2.1), one easily verifies conditions (O1) and (O3) of Proposition 2.5 hold. Condition (O2) is the same as

\[
\sum_{i=1}^{k^*} \left( c_i^2 \| z_0 - z_i \|_2^2 + z_i^\top Q_i z_i \right) + \sum_{j=k^*+1}^{\ell} z_j^\top (Q_j - c_j^2 I_4) z_j \geq 0, \quad \forall z_i \in \mathbb{R}^4,
\]

which holds true because \( Q_i \) are positive semidefinite as per Lemma 2.1 and \( Q_j \geq c_j^2 I_4 \).

For the second part, assume \( c_j^2 > (w_0^*)^\top Q_j w_0^* \), and we will show \( W^* (w^*)^\top \) does not minimize (SDR). For this we need to show that conditions (O1), (O2), and (O3) of Proposition 2.5 can not be simultaneously satisfied by \( w_0^* \) and any \( \hat{\mu} \) and any \( \mathcal{D} \), where \( \mathcal{D} \) is of the form (2.10).
Assume (O1) and (O3) are true and we will prove (O2) is false. In particular, (O3) implies \( \hat{\mu} = -\sum_{i=1}^{k^*} c_i^2 \). And (O1) implies \( 2[\mathcal{D}]_{\ell \ell} + c_i^2 I_4 - Q_j w_0^* = 0 \), which yields the inequality
\[
2(w_0^*)^\top [\mathcal{D}]_{\ell \ell} w_0^* = (w_0^*)^\top Q_j w_0^* - c_i^2 < 0.
\]
Hence \( 2[\mathcal{D}]_{\ell \ell} \) must have a negative eigenvalue. Let \( z = [z_0; \ldots; z_\ell] \in \mathbb{R}^{(\ell+1)} \) be such that \( z_i = 0 \) for any \( i \neq \ell \) and \( z_\ell \in \mathbb{S}^3 \) is an eigenvector of \( 2[\mathcal{D}]_{\ell \ell} \) corresponding to that negative eigenvalue. Then \( z^\top (Q - \hat{\mu} B - \mathcal{D}) z = 2z_\ell^\top [\mathcal{D}]_{\ell \ell} z_\ell < 0 \), and condition (O2) is not true. This finishes the proof of the second statement.

**Remark 3.3 (Noiseless Point Sets with Random Outliers).** If outlier \( (y_j, x_j) \) is randomly drawn from \( \mathbb{R}^3 \times \mathbb{R}^3 \) according to some continuous probability distribution, then with probability 1 we have \( \|y_j\|_2 \neq \|x_j\|_2 \) (Lemma 2 of [54]), which implies \( \lambda_{\min}(Q_j) > 0 \). Thus, (SDR) is always tight to such random outliers, if \( c_i^2 \to 0 \). Note that this discussion is theoretical and does not apply to the case where \( \|y_j\|_2 = 0 \) is nonzero but is below machine accuracy, as \( c_i^2 \) can not be set even smaller (we only consider \( c_i^2 > 0 \)).

If the condition \( c_j^2 < \lambda_{\min}(Q_j) \) of the first statement in Theorem 3.2 holds then \( Q_j \) will always be rejected by (TLS-Q) as an outlier. In fact, since \( \lambda_{\min}(Q_j) \) can be easily computed (Lemma 2.1), in practice one usually throws away the point pairs \( (y_j, x_j) \)'s for which \( c_j^2 < \lambda_{\min}(Q_j) \) as a means of preprocessing (cf. [12, 47]), and these point pairs might not enter into the semidefinite optimization. Our result here thus shows that (SDR) can distinguish this type of “simple” outliers, as long as \( c_i^2 \) is chosen sufficiently small.

In the second statement of Theorem 3.2, if \( c_j^2 > (w_0^*)^\top Q_j w_0^* \) holds true for outlier \( Q_j \), then (TLS-Q) would attempt to minimize \( w_0^* Q_j w_0^* + \sum_{i=1}^{k^*} w_0^* Q_i w_0^* \) over \( w_0^* \in \mathbb{S}^3 \) at least—an outlier showed up in the eigenvalue optimization—thus the global minimizer of (TLS-Q) is unlikely to be \( w_0^* \). This is why we do not expect \( \mathcal{W}^* \) to globally minimize (SDR).

Admittedly, Theorem 3.2 leaves a gap: What can we say about the tightness of (SDR) if
\[
\lambda_{\min}(Q_j) < c_j^2 < (w_0^*)^\top Q_j w_0^* ?
\]

While our empirical observation suggests that \( \mathcal{W}^* \) does not globally minimize (SDR) if (3.2) holds (with \( k^* < \ell \)), the analysis of this case without further assumptions on outliers appears hard. The difficulty is that the outliers \( Q_j \)'s can be adversarial that \( (w_0^*)^\top Q_j w_0^* \) is arbitrarily close\(^3\) to 0 while \( \lambda_{\min}(Q_j) = 0 \) for every \( j > k^* \). Thus, the value of Theorem 3.2 is in that it shows that (SDR) can only handle “simple” outliers that can be filtered out, and thus reveals a fundamental limit on the performance of (SDR).

Next, we consider the situation where outliers \( Q_j \)'s can not be simply removed by preprocessing, e.g., \( \lambda_{\min}(Q_j) = 0 \). In particular, we assume the outliers are clustered and show that the (SDR) under investigation is even more vulnerable:

**Theorem 3.4 (Noiseless Point Sets with Clustered Outliers).** With the notation of Theorem 3.2, further suppose outliers \( Q_{k^*+1}, \ldots, Q_{\ell} \) are “clustered” in the sense that
\[
Q_{k^*+1} w_0^{cl} = \cdots = Q_{\ell} w_0^{cl} = 0
\]

\(^3\)Alternatively, if \( (w_0^*)^\top Q_j w_0^* \) is small, then \( Q_j \) might be treated as noisy data rather than an outlier. We consider such noisy case in subsection 3.3 (without outliers) and subsection 3.4 (with outliers).
with \( w_0^{cl} \in S^3 \) some unit quaternion that is different from \( \pm w_0^* \). If

\[
1 - \frac{\sum_{j=k^*+1}^{\ell} c_j^2}{2 \sum_{i=1}^{k^*} c_i^2} < \| (w_0^{cl})^\top w_0^* \|,
\]

then \( \mathcal{W}^* \) does not globally minimize (SDR).

**Proof of Theorem 3.4.** Here we follow the proof idea in Theorem 3.2. Let \( \hat{\mu} := -\sum_{i=1}^{k^*} c_i^2 \), \( \hat{\mathcal{D}} \) of the form (2.10), and they satisfy (O1) (and (O3)) with \( w_0^* \). We will show that, given the clustered outliers which satisfy (3.3) and (3.4), the dual feasibility condition (O2) is not true. Suppose without loss of generality \( (w_0^{cl})^\top w_0^* > 0 \). With \( z_1 = \cdots = z_{k^*} = w_0^1 \), and \( z_{k^*+1} = \cdots = z_{\ell} = z_0 = w_0^0 \), (O2) becomes

\[
\sum_{i=1}^{k^*} c_i^2 (\| w_0^{cl} \|_2^2 + \| w_0^* \|_2^2 - 2 (w_0^{cl})^\top w_0^*) + \sum_{j=k^*+1}^{\ell} (w_0^{cl})^\top (Q_j - c_j^2 I_4) w_0^0 \geq 0
\]

\[
\Leftrightarrow 1 - \frac{\sum_{j=k^*+1}^{\ell} c_j^2}{2 \sum_{i=1}^{k^*} c_i^2} \geq (w_0^{cl})^\top w_0^*,
\]

which violates assumption (3.4). This finishes the proof.

The *clustered* outliers of Theorem 3.4 defined in the sense of (3.3) mean that the outlier pairs \( (y_j, x_j) (j > k^*) \) are related by the same 3D rotation \( R_0^0 \) that correspond to \( w_0^{cl} \), that is \( y_j = R_0^0 x_j \), \( \forall j > k^* \) (Lemma 2.1). Clustered outliers can be thought of as a special type of adversarial outliers, the latter usually used to study the robustness of algorithms in the worst case; it should be distinguished from data clustering [33, 23].

To understand condition (3.4) of Theorem 3.4, consider a situation where all truncation parameters are equal, \( c_1^2 = \cdots = c_{k^*}^2 \). Then (3.4) simplifies to \( 1 - (\ell - k^*)/(2k^*) < |(w_0^{cl})^\top w_0^*| \); also note that \( |(w_0^{cl})^\top w_0^*| \in [0, 1) \). Thus, if \( \ell - k^* > 2k^* \), then (3.4) always holds, and so \( \mathcal{W}^* \) never globally minimizes (SDR), which is forgivable as in this case \( w_0^* \) neither globally minimizes (TLS-Q). However, even if the number of outliers is only half the number of inliers, i.e., \( \ell - k^* = k^*/2 \), Theorem 3.4 implies that \( \mathcal{W}^* \) would still fail to globally minimize (SDR) as long as \( |(w_0^{cl})^\top w_0^*| > 3/4 \), but \( w_0^0 \) would in general globally minimize (TLS-Q) with suitable \( c_j^2 \) (cf. [61]). In this sense, (SDR) is strictly less robust to outliers than (TLS-Q).

Finally, we note that Theorem 3.4 might be overly pessimistic. In fact, experiments show that (SDR) is robust to 40%-50% outliers \( (y_j, x_j) \)'s, where \( y_j \) and \( x_j \) are sampled uniformly at random from \( S^2 \) (so \( \lambda_{\min}(Q_j) = 0 \) by Lemma 2.1). Two factors account for this empirically better behavior: i) Such random outliers are less adversarial than clustered ones, ii) the extra projection step that converts the global minimizer of (SDR) to a unit quaternion alleviates to some extent the issue of \( \mathcal{W}^* \) not minimizing (SDR). In retrospect, there are two downsides in our analysis of Theorems 3.2 and 3.4: (1) We have not taken such extra projection step into account, and (2) we only showed that \( \mathcal{W}^* \) might not minimize (SDR) but have not proved how far the global minimizers of (SDR) can be from \( \mathcal{W}^* \).

### 3.3. The Noisy + Outlier-Free Case

The noisy case, even without outliers, is more difficult to penetrate than previous cases. A general reason for this is that the global minimizers of (QCQP) and (SDR) are now complicated functions of noise. One might wonder
whether Theorem 3.1 can be extended to the noisy + outlier-free case using some continuity argument. In fact, [21] shows that, under certain conditions, if the noiseless version of the Schor relaxation of some QCQP is tight, then its noisy version is also tight. While this result is quite general, its conditions are abstract and, at least for non-experts, hard to verify. To our case, [21] is not directly applicable, as in our problem the truncation parameters $c_i^2$ also have impacts on the tightness, and the approach of [21] does not model, and thus could not control the values of $c_i^2$. Instead, our analysis must take both $c_i^2$ and noise into account.

We begin our analysis by decomposing the corrupted data matrix $Q_i$ of (TLS-Q) into the pure data part $P_i$ and noise part $E_i + \|\epsilon_i\|^2 I_4$:

**Lemma 3.5.** Let $R_0 \in \text{SO}(3)$. If $(y_i, x_i)$ is an inlier that obeys (1.1), then we have

$$
\|y_i - R_0 x_i\|^2 = \|\epsilon_i + R_0^* x_i - R_0 x_i\|^2 = \|P_i x_i - R x_i\|^2 + \|\epsilon_i\|^2 + 2\epsilon_i^T (R_0^* x_i - R_0 x_i)
$$

where $w_0 \in S^3$ is the unit quaternion representation of $R_0$, and $P_i$ and $E_i$ are $4 \times 4$ symmetric matrices that respectively satisfy the following properties:

- $P_i$ is positive semidefinite with its entries depending on $y_i$ and $x_i$, and it has two different eigenvalues $4\|x_i\|^2$ and 0, each of multiplicity 2. The ground-truth unit quaternion $w_0^*$ is an eigenvector of $P_i$ corresponding to eigenvalue 0, i.e., $P_i w_0^* = 0$. In particular, we have $Q_i w_0^* = P_i w_0^* = 0$ in the noiseless case.

- $E_i$ has entries depending on $y_i$, $x_i$, and noise $\epsilon_i$, and has two different eigenvalues $2 \epsilon_i^T R_0^* x_i + 2\|\epsilon_i\|^2\|x_i\|^2$ and $2 \epsilon_i^T R_0^* x_i - 2\|\epsilon_i\|^2\|x_i\|^2$, each of multiplicity 2. We have $w_0^* E_i w_0^* = 2\epsilon_i^T (R_0^* x_i - R_0 x_i)$ and in particular $(w_0^*)^T E_i w_0^* = 0$.

**Proof of Lemma 3.5.** We start with the following equality:

$$
\|y_i - R_0 x_i\|^2 = \|\epsilon_i + R_0^* x_i - R_0 x_i\|^2 = \|R_0^* x_i - R x_i\|^2 + \|\epsilon_i\|^2 + 2\epsilon_i^T (R_0^* x_i - R_0 x_i)
$$

From the proof of Lemma 2.1 and in particular (2.3), we see that the first term $\|R_0^* x_i - R_0 x_i\|^2$ is equal to $w_0^T P_i w_0$, where $P_i := Q_i(R_0^* x_i, x_i)$ is a $4 \times 4$ positive semidefinite matrix with eigenvalues $4\|x_i\|^2$, $4\|x_i\|^2$, 0, 0, and $w_0$ is the unit quaternion that represents $R_0$. In particular, we have $(w_0^*)^T P_i w_0^* = \|R_0^* x_i - R_0 x_i\|^2 = 0$, which implies $P_i w_0^* = 0$.

On the other hand, from the proof of Lemma 2.1 and in particular (2.4), we see the equality $-2\epsilon_i^T R_0 x_i = -2w_0^T U(\epsilon_i, x_i) w_0$ where $-2U(\epsilon_i, x_i)$ is symmetric and has two different eigenvalues $2\|\epsilon_i\|^2\|x_i\|^2$ and $-2\|\epsilon_i\|^2\|x_i\|^2$, each of multiplicity 2. Then $E_i := 2(\epsilon_i^T R_0^* x_i) I_4 - 2U(\epsilon_i, x_i)$ satisfies the claimed properties.

Our investigation into the noisy case is tightly related to the eigengap $\zeta$, defined as the ratio between the second smallest eigenvalue of $\sum_{i=1}^\ell Q_i$ and its minimum eigenvalue:

$$
\zeta := \frac{\lambda_{\text{min}}(\sum_{i=1}^\ell Q_i)}{\lambda_{\text{min}}(\sum_{i=1}^\ell Q_i)}
$$

In analysis of eigenvalue algorithms (cf. [18]), the eigengap is typically defined as the difference between two consecutive eigenvalues of some matrix. Our eigengap (3.6) that takes the division of two smallest eigenvalues is not standard, but it will be convenient for our purpose. Also note that, $Q_i$ is a perturbed version of $P_i$ by noise $\epsilon_i$ (Lemma 3.5), so $\lambda_{\text{min}}(\sum_{i=1}^\ell Q_i)$ is in
general nonzero, while it indeed approaches zero if $\epsilon_i \to 0$. Clearly $\zeta \geq 1$. Moreover, we have the following immediate observation:

**Remark 3.6.** If (TLS-Q) has a unique solution and if $c_i^2 > \hat{w}_i^\top Q_i \hat{w}_0 (\forall i)$, then $\zeta > 1$.

We are now ready to state the following result:

**Theorem 3.7 (Noisy and Outlier-Free Point Sets).** Consider (TLS-Q) with noisy inliers $\{Q_i\}_{i=1}^\ell$ and $\hat{w}_0 \in S^3$ its global minimizer. Let $\hat{w} := [\hat{w}_0; \ldots; \hat{w}_0] \in \mathbb{R}^{4(\ell+1)}$. Suppose $\zeta \geq \ell/(\ell - 1)$. (SDR) is tight as long as

$$c_i^2 > \hat{w}_0^\top Q_i \hat{w}_0 + \|Q_i \hat{w}_0\|_2^2 + \frac{|d_i| + d_i}{2}, \quad \forall i = 1, \ldots, \ell$$

(3.7)

with $d_i := \sum_{i=1}^\ell \hat{w}_0^\top Q_i \hat{w}_0 - \hat{w}_0^\top Q_i \hat{w}_0 + \frac{\lambda_{\text{max}}}{\zeta(\ell - 1)} \left( \sum_{j \neq i} (Q_i - Q_j) \right)$.

Moreover, the angle $\tau^*_{\hat{w}_0}$ between $\hat{w}_0$ and the ground-truth unit quaternion $w_0^* \in S^3$ grows proportionally with the magnitude of noise $\epsilon_i$:

$$\sin^2(\tau^*_{\hat{w}_0}) \leq \frac{4\sum_{i=1}^\ell \|\epsilon_i\|_2^2 \|x_i\|_2^2}{\lambda_{\text{min}}^2(\sum_{i=1}^\ell P_{\epsilon_i})}, \quad \sin^2(\tau^*_{\hat{w}_0}) := 1 - (\hat{w}_0^\top w_0^*)^2$$

In (3.8), each $P_{\epsilon_i}$ corresponds to the “pure data” part of $Q_i$ that satisfies $P_{\epsilon_i} w_0^* = 0, P_{\epsilon_i} \geq 0$ (Lemma 3.5), and $\lambda_{\text{min}}^2(\cdot)$ denotes the second smallest eigenvalue of a matrix.

**Proof of Theorem 3.7.** Since (3.7) implies $c_i^2 > \hat{w}_0^\top Q_i \hat{w}_0$, (TLS-Q) preserves all inliers with minimum $\sum_{i=1}^\ell \hat{w}_0^\top Q_i \hat{w}_0$. So $\hat{\mu} := \sum_{i=1}^\ell (\hat{w}_0^\top Q_i \hat{w}_0 - c_i^2)$ satisfies (O3) of Proposition 2.5.

Let $\hat{V} := [\hat{V}_0; \hat{w}_0] \in \mathbb{R}^{4 \times 4}$ form an orthonormal basis of $\mathbb{R}^4$; $\hat{V}_0 \in \mathbb{R}^{4 \times 3}$ satisfies $\hat{V}_0^\top \hat{w}_0 = 0$ and $\hat{V}_0^\top \hat{V}_0 = I_3$. Let $\hat{D}$ be the form (2.10) with each $[\hat{D}]_{0i}$ satisfying

$$[\hat{D}]_{0i} := \hat{S} - \frac{1}{2} (Q_i - c_i^2 I_4), \quad \hat{S} := \hat{V} \begin{bmatrix} \hat{T}_i & 0 \\ 0 & 0 \end{bmatrix} \hat{V}^\top$$

(3.9)

where $\hat{T}_i$ is a $3 \times 3$ symmetric matrix defined as

$$\hat{T}_i := \frac{\zeta - \ell}{\zeta} \hat{V}_0^\top Q_i \hat{V}_0 + \sum_{j=1}^\ell \hat{V}_0^\top Q_j \hat{V}_0 - \frac{\left( \sum_{i=1}^\ell \hat{w}_0^\top Q_j \hat{w}_0 \right)}{\zeta(\ell - 1)} I_3.$$  

(3.10)

Clearly $\hat{S}_i \hat{w}_0 = 0$, with which one verifies that (O1) of Proposition 2.5 is fulfilled.

It remains to prove (O2) holds true. For this we need the following lemma. Lemma 3.8 is a bit technical, so we put its proof into the appendix.

**Lemma 3.8.** Let $[\hat{D}]_{0i}, \hat{S}_i$, and $\hat{T}_i$ be defined in (3.9) and (3.10) respectively. Suppose that the assumptions of Theorem 3.7, namely $\zeta \geq \ell/(\ell - 1)$ and (3.7), hold. Then

$$\sum_{i=1}^\ell \hat{S}_i = \sum_{i=1}^\ell (Q_i - \hat{w}_0^\top Q_i \hat{w}_0 I_4)$$

$$\hat{S}_i \succeq 0, \quad \forall i = 1, \ldots, \ell$$

$$\hat{S}_i + c_i^2 I_4 - Q_i > 0, \quad \forall i = 1, \ldots, \ell$$
With Lemma 3.8 and definitions of \( |\mathcal{D}|_0 \) (3.9), we can now write (O2) as \((\forall z_i \in \mathbb{R}^4)\)

\[
\sum_{i=1}^{\ell} \left( z_i^0 \left( c_i^2 - w_0^T Q_i w_0 \right) z_i + z_i^T \left( 2 \tilde{S}_i + c_i^2 I_4 - Q_i \right) z_i - 2 z_i^0 \left( \tilde{S}_i + c_i^2 I_4 - Q_i \right) z_i \right) \geq 0
\]

\[
\Leftrightarrow \sum_{i=1}^{\ell} \left( z_i^0 \left( \tilde{S}_i + c_i^2 I_4 - Q_i \right) z_i + z_i^T \left( 2 \tilde{S}_i + c_i^2 I_4 - Q_i \right) z_i - 2 z_i^0 \left( \tilde{S}_i + c_i^2 I_4 - Q_i \right) z_i \right) \geq 0
\]

\[
\Leftrightarrow \sum_{i=1}^{\ell} (z_i - z_0)^T \left( \tilde{S}_i + c_i^2 I_4 - Q_i \right) (z_i - z_0) + \sum_{i=1}^{\ell} z_i^T \tilde{S}_i z_i \geq 0
\]

which holds, as \( \tilde{S}_i + c_i^2 I_4 - Q_i \) and \( \tilde{S}_i \) are both positive semidefinite for every \( i = 1, \ldots, \ell \) (Lemma 3.8). We have thus verified all conditions of Proposition 2.5, thus (SDR) is tight.

It now remains to prove the error bound (3.8). With Lemma 3.5 we write \( Q_i = P_i + E_i + \| e_i \|^2 I_4 \), and we know that \( \hat{w}_0 \) is an eigenvector of \( \sum_{i=1}^{\ell} (P_i + E_i) \) corresponding to its minimum eigenvalue. Hence, with the optimality of \( \hat{w}_0 \) and Lemma 3.5, we have

\[
\sum_{i=1}^{\ell} \hat{w}_0^T (P_i + E_i) \hat{w}_0 \leq \sum_{i=1}^{\ell} (w_0^i)^T (P_i + E_i) w_0^i = 0
\]

\[
\Rightarrow \sum_{i=1}^{\ell} \hat{w}_0^T P_i \hat{w}_0 \leq - \sum_{i=1}^{\ell} w_0^i E_i w_0^i = - \sum_{i=1}^{\ell} 2 e_i^T (R_0^i - \hat{R}_0) x_i \leq 4 \sum_{i=1}^{\ell} \| e_i \|_2 \| x_i \|_2
\]

Here we recall that \( \hat{R}_0 \) is the global minimizer of (TLS-R) that corresponds to \( \hat{w}_0 \), and used Lemma 3.5. Since \( \sum_{i=1}^{\ell} P_i \) is positive semidefinite with \( P_i w_0^i = 0 \), from the eigen decomposition of \( \sum_{i=1}^{\ell} P_i \) we can easily obtain the inequality

\[
(1 - (w_0^T w_0^i)^2) \lambda_{\text{min}2} \left( \sum_{i=1}^{\ell} P_i \right) \leq \sum_{i=1}^{\ell} w_0^T P_i w_0,
\]

where \( \lambda_{\text{min}2}(\cdot) \) denotes the second smallest eigenvalue of a matrix. The proof is complete. \( \blacksquare \)

Theorem 3.7 is better understood via numerics. We take randomly generated \( \ell = 100 \) point pairs \( (y_i, x_i) \)’s with \( x_i \sim \mathcal{N}(0, I_3) \), and add different levels of Gaussian noise \( e_i \sim \mathcal{N}(0, \sigma^2 I_3) \), where \( \sigma \) ranges from 1\% to 10\%. The values of \( \lambda_{\text{min}}(\sum_{i=1}^{\ell} Q_i) \) and \( \lambda_{\text{min}2}(\sum_{i=1}^{\ell} Q_i) \), and thus \( \zeta \), are shown in Figure 1a, where one might observe that \( \zeta \approx 250 \) for 10\% noise, \( \zeta \approx 2500 \) for 1\% noise, and, in general, \( \zeta = \infty \) for the noiseless case. This empirically validates the assumption \( \zeta \geq \ell / (\ell - 1) = 100/99 \).

We then elaborate the more complicated condition (3.7). First we recall that \( c_i^2 > w_0^T Q_i w_0 \) is essential for (TLS-Q) to preserve all inliers. Second, we argue that the term \( \| Q_i \|_2 \) in (3.7) is also essential, as it accounts for the fact that noise destroys the inequality \( \lambda_{\text{min}}(Q_i) - w_0^T Q_i w_0 \geq 0 \) which holds in the noiseless case (where \( \lambda_{\text{min}}(Q_i) = w_0^T Q_i w_0 = 0 \)) but gets violated (in general) in the presence of noise. Finally, (3.7) also incurs a curious term \((\| d_i \| + d_i) / 2\), with \( d_i \) defined in a sophisticated way (3.7). If \( d_i < 0 \) then this term is 0. Thus it remains to understand the values of \( |d_i| \). In particular, we plotted the values of \( |d_i| \) in Figure
for details. In comparison to \( \hat{Q}_i \hat{w}_0 \to 0 \) and (in general) \( \zeta \to \infty \), hence \( d_i \to 0 \) by definition (3.7). Overall, condition (3.7) degenerates into \( c_i^2 > 0 \) in the noiseless case.

While the term \((d_i + d_j)/2\) is quite small (Figure 1b) and sometimes harmless (e.g., when \( d_i < 0 \)), it appears as an artifact of our analysis, and we expect an ideal condition for the noisy + outlier-free case to be \( c_i^2 > \| \hat{Q}_i \hat{w}_0 \|_2 \). However, proof under this alternative condition demands showing some matrix inequality that involves a sum of matrix inverses always holds; we were not able to prove it.

Finally, we discuss the error bound (3.8). It becomes zero as \( \epsilon_i \to 0 \) and thus \( \hat{w}_0 = w_0^* \), provided that \( \| x_i \|_2/\lambda_{\min 2}(\sum_{i=1}^\ell P_i) \) is not too large. The denominator \( \lambda_{\min 2}(\sum_{i=1}^\ell P_i) \) seems inevitable, as it usually determines the stability of solving a minimum eigenvalue problem: If \( \lambda_{\min}((\sum_{i=1}^\ell P_i)) = 0 \), then \( \hat{w}_0 \) can be arbitrarily far from \( w_0^* \) even in the slightest presence of noise. That said, \( \| x_i \|_2/\lambda_{\min 2}(\sum_{i=1}^\ell P_i) \) is actually quite small and well-behaved, at least for random Gaussian data; see section 4 for details.

3.4. The Noisy + Outliers Case. With the proof ideas of Theorems 3.2 and 3.7, we get:

\[ \text{Theorem 3.9 (Noisy Point Sets with Outliers).} \quad \text{Let} \ Q_1, \ldots, Q_{k^*} \ \text{be inliers, the rest} \ Q_j \ \text{'s outliers, and} \ \hat{w}_0 \ \text{a global minimizer of} \ \langle \text{TLS-Q} \rangle. \ \text{Define} \]

\[ c_{\text{in}} := \frac{\lambda_{\min 2}(\sum_{i=1}^{k^*} Q_i)}{\lambda_{\min}((\sum_{i=1}^{k^*} Q_i))} \]

Assume (1) \( c_{\text{in}} \geq k^*/(k^* - 1) \), (2) for every \( j = k^* + 1, \ldots, \ell \), we have \( 0 < c_j^2 < \lambda_{\min}(Q_j) \), (3) for every \( i = 1, \ldots, k^* \), (3.7) holds with \( d_i \) now defined as

\[ d_i := \frac{\sum_{j=1}^{k^*} \hat{w}_0^\top Q_j \hat{w}_0}{k^*} - \hat{w}_0^\top Q_i \hat{w}_0 + \frac{\lambda_{\max}(\sum_{j \neq i} (Q_i - Q_j))}{\zeta(k^* - 1)}. \]

Then (SDR) is tight and, similarly to (3.8) we have

\[ \sin^2(\hat{\alpha}_i^2) \leq \frac{4 \sum_{i=1}^{k^*} \| \epsilon_i \|_2^2 \| x_i \|_2^2}{\lambda_{\min 2}(\sum_{i=1}^{k^*} P_i)}, \quad \sin^2(\hat{\alpha}_i^2) := 1 - (\hat{w}_0^\top \hat{w}_0^*)^2. \]
Here we recall that $w_0^* \in S^3$ is the ground-truth unit quaternion, and each $P_i$ is the “pure data” part of $Q_i$ that satisfies $P_i w_0^* = 0, P_i \succeq 0$ (Lemma 3.5).

Proof. The given assumptions ensure that (TLS-Q) rejects all outliers and admit all inliers, and the minimum of (TLS-Q) is $\sum_{i=1}^{k^*} \hat{w}_0^* Q_i \hat{w}_0 + \sum_{j=k^*+1}^{\ell} c^2_j$. For $i = 1, \ldots, k^*$ let $x_0^i$ be defined as in (3.9), and for $j = k^* + 1, \ldots, \ell$ let $[\hat{D}]_{0j} := \frac{1}{2}(Q_j - c^2_j I_4)$. Let $\hat{\mu} := \sum_{i=1}^{k^*} (\hat{w}_0^* Q_i \hat{w}_0 - c^2_i)$ One then verifies the optimality conditions (O1) and (O3) of Proposition 2.5 are satisfied. (O2) is equivalent to $(\forall \mathbf{x}_i \in \mathbb{R}^4)$

$$
-\hat{\mu} \|z_0\|^2_2 + \sum_{i=1}^{k^*} \left( 2z_0^\top [\hat{D}]_{0i} z_i - z_0^\top (2[\hat{D}]_{0i} - Q_i + c^2_i I_4) z_i \right) + \sum_{i=k^*+1}^{\ell} z_j(Q_j - c^2_j I_4) z_j \geq 0,
$$

Since $Q_j \succeq c^2_j I_4$, the outlier term is non-negative. Under the given assumptions, one can replace $\ell$ by $k^*$ in the proof of Theorem 3.7 and then find the inlier term is also non-negative. This finishes proving (O2) and thus the tightness of (SDR). The error bound (3.13) follows from the proof of Theorem 3.7 with $\ell$ also replaced by $k^*$.

4. Probabilistic Interpretation of Error Bound (3.8). Since the entries of $P_i$ depend quadratically on entries of $\mathbf{x}_i$, we naturally expect $\lambda_{\min2}(\sum_{i=1}^{\ell} P_i) = \Theta(\sum_{i=1}^{\ell} \|x_i\|^2_2)$ for random points $\mathbf{x}_i$'s; Proposition 4.1 confirms that this is indeed true with high probability:

Proposition 4.1. Suppose that every 3D point $\mathbf{x}_i$ has $N(0, 1)$ entries. Let $t$ be some positive constant and let $\ell$ be large enough in the sense that

$$
\ell \geq (4 + 2\sqrt{3} + 2t) \sqrt{\ell} + (\sqrt{3} + t)^2.
$$

Then the following holds with probability at least $1 - \exp(-t^2/2) - 2\exp(-3t^2/8)$:

$$
\frac{1}{4} \leq \frac{\sum_{i=1}^{\ell} \|x_i\|^2_2}{\lambda_{\min2}(\sum_{i=1}^{\ell} P_i)} \leq \frac{3}{4}
$$

Recall that $P_i$ is the pure data part of $Q_i$ in the noisy case (cf. Lemma 3.5).
**Remark 4.2 (Asymptotic Behavior).** With Lemmas 4.3 to 4.5 and [55], we get

\[
\ell \to \infty \Rightarrow \frac{\sum_{i=1}^\ell \|x_i\|^2_2}{\lambda_{\min}^2(\sum_{i=1}^\ell P_i)} = \frac{3}{8}
\]

**Proof of Proposition 4.1.** Since \(x_i\)'s have i.i.d. \(N(0, 1)\) entries, \(\lambda_{\min}^2(\sum_{i=1}^\ell P_i)\) is not equal to zero with probability 1. Then Lemma 4.5 implies that the left bound holds. For the right bound, applying the union bound with Lemmas 4.3 to 4.5, it holds that

\[
\sum_{i=1}^\ell \|x_i\|^2_2 \leq \frac{3\ell + 3\sqrt{\ell}}{4 \cdot (2\ell - 3\sqrt{\ell} - 2(\sqrt{3} + t)\sqrt{\ell} - (\sqrt{3} + t)^2)}
\]

with probability at least \(1 - \exp(-t^2/2) - 2\exp(-3t^2/8)\). Next, from (4.1) we get

\[
2\ell - 3\sqrt{\ell} - 2(\sqrt{3} + t)\sqrt{\ell} - (\sqrt{3} + t)^2 \geq \ell + \sqrt{\ell},
\]

which means the right-hand side of (4.4) does not exceed \(3/4\). The proof is finished. \(\blacksquare\)

**Lemma 4.3 (Example 2.11 of [55]).** If each \(x_i\) in \(\mathbb{R}^3\) has i.i.d. \(N(0, 1)\) entries, then it holds for some \(t > 0\) with probability at least \(1 - \exp(-t^2/2)\) that

\[
3\ell - 3\sqrt{\ell} \leq \sum_{i=1}^\ell \|x_i\|^2_2 \leq 3\ell + 3\sqrt{\ell}, \quad t > 0
\]

**Lemma 4.4 (Theorem 6.1 and Example 6.2 of [55]).** If each \(x_i\) in \(\mathbb{R}^3\) has i.i.d. \(N(0, 1)\) entries, then for some \(t > 0\) with probability at least \(1 - \exp(-t^2/2)\) we have

\[
\lambda_{\max}(\sum_{i=1}^\ell x_i x_i^\top) \leq \ell + 2(\sqrt{3} + t)\sqrt{\ell} + (\sqrt{3} + t)^2
\]

**Lemma 4.5.** Recall that \(P_i\) is the pure data part of \(Q_i\) in the noisy case (cf. Lemma 3.5), and that \(\lambda_{\min}^2(\cdot)\) denotes the second smallest eigenvalue of a matrix. We have

\[
\lambda_{\min}^2(\sum_{i=1}^\ell P_i) = \sum_{i=1}^\ell 4\|x_i\|^2_2 - \lambda_{\max}(\sum_{i=1}^\ell 4x_i x_i^\top)
\]

**Proof of Lemma 4.5.** With any \(w_2 \in S^3\), \(\phi\) and \(b\) the rotation angle and rotation axis of \(R_2^\top R_0^\top\) respectively, and with Lemma 4.6, we have the following equivalence:

\[
\begin{align*}
\mathbf{w}_2 \mathbf{w}_0^\top = 0 & \iff \phi = \pi \iff R_2^\top R_0^\top = 2bb^\top - I_3
\end{align*}
\]

In the last equivalence we used the axis-angle representation of 3D rotations

\[
R_2^\top R_0^\top = bb^\top + [b]_\times \sin(\phi) + (I_3 - bb^\top) \cos(\phi)
\]
and substituted $\phi = \pi$ into it; here, $[b]_\times \in \mathbb{R}^{3 \times 3}$ is the cross product matrix that satisfies $[b]_\times q = b \times q$ for any $q \in \mathbb{R}^3$. On the other hand, Lemma 2.1 implies

$$w_2^\top \left( \sum_{i=1}^\ell P_i \right) w_2 = \sum_{i=1}^\ell \| R_0^i x_i - R_0^i x_i \|_2^2 = \sum_{i=1}^\ell 2 \| x_i \|_2^2 - \sum_{i=1}^\ell 2 \langle x_i, R_2^i R_0^i x_i \rangle.$$

Clearly, as the unit quaternion $w_2$ varies in $\mathbb{S}^3$ with $w_2^\top w_0^* = 0$, the axis of $b$ of $R_2^i R_0^*$ can be any element of $\mathbb{S}^2$, and vice versa. Combine the above, and we arrive at

$$\min_{w \in \mathbb{S}^3} w_2^\top \left( \sum_{i=1}^\ell P_i \right) w_2 = \sum_{i=1}^\ell 2 \| x_i \|_2^2 - \max_{b \in \mathbb{S}^2} \sum_{i=1}^\ell 2 b_i^\top (2b_i^\top I_3) x_i$$

Since $w_0^*$ corresponds to the minimum eigenvalue 0 of $\sum_{i=1}^\ell P_i$, we know that the left-hand side of (4.7) is equal to $\lambda_{\min2} \left( \sum_{i=1}^\ell P_i \right)$. Also observe that the right-hand side of (4.7) is equal to that of (4.6), and we have finished the proof.

**Lemma 4.6.** For two 3D rotations $R_1$ and $R_2$, their respective unit quaternion representations $w_1$ and $w_2$, and the rotation angle $\phi_{12} \in [0, \pi]$ of $R_1^\top R_2$, we have

$$w_1^\top w_2 = \cos \left( \frac{\phi_{12}}{2} \right).$$

In particular, we get that $w_1^\top w_2 = 0$ if and only if $\phi_{12} = \pi$.

**Proof of Lemma 4.6.** Denote by $b_1$ (resp. $b_2$) and $\phi_1$ (resp. $\phi_2$) the rotation axis and angle of $R_1$ (resp. $R_2$). Then, from a well-known relation between unit quaternions and axis-angle representation of rotations, we know that $w_1 = [\cos(\phi_1/2) ; b_1 \sin(\phi_1/2)]$ and $w_2 = [\cos(\phi_2/2) ; b_2 \sin(\phi_2/2)]$, and therefore we have

$$w_1^\top w_2 = \cos \left( \frac{\phi_1}{2} \right) \cdot \cos \left( \frac{\phi_2}{2} \right) + b_1^\top b_2 \cdot \sin \left( \frac{\phi_1}{2} \right) \cdot \sin \left( \frac{\phi_2}{2} \right).$$

Since $R_1^\top$ is a 3D rotation with angle $\phi_1$ and axis $-b_1$, it is well known that the composition $R_1^\top R_2$ of two 3D rotations has its rotation angle $\phi_{12}$ satisfying

$$\cos \left( \frac{\phi_{12}}{2} \right) = \cos \left( \frac{\phi_1}{2} \right) \cdot \cos \left( \frac{\phi_2}{2} \right) - (-b_1)^\top b_2 \cdot \sin \left( \frac{\phi_1}{2} \right) \cdot \sin \left( \frac{\phi_2}{2} \right).$$

Equality (4.8) can be proved via spherical trigonometry, or (1.2), or Rodrigues’ rotation formula, with some algebraic manipulations. We have thus finished the proof.

**5. Discussion and Future Work.** We have investigated the tightness of a semidefinite relaxation (SDR) of truncated least-squares for robust rotation search in four different cases, and in each case we either showed improvements over prior work or proved new theoretical results. Our investigation can potentially be borrowed to understand semidefinite relaxations of many other geometric vision tasks; see [60] for 6 examples of truncated least-squares and see also [21, 3, 11].
As is common in the optimization literature, the relaxation we analyzed is at the first (i.e., lowest) relaxation order of the Lasserre hierarchy [37], or otherwise known as the Shor relaxation [51]. A tighter relaxation that has quadratically more constraints than (SDR) exists (cf. [59]). However, analyzing this tighter relaxation is significantly harder, as one needs to either (1) construct quadratically more dual certificates during the proof, or (2) use more abstract optimality conditions (cf. [20, 21]). Therefore, we leave this challenging question to future work.

**Appendix A. Proof of Lemma 3.8.**

**Proof of Lemma 3.8.** We first prove the equality \( \sum_{i=1}^{\ell} \hat{S}_i = \sum_{i=1}^{\ell} (Q_i - w_0^T Q_i w_0 I_4) \). First, by definition of \( \hat{T}_i \) (3.10) and simple calculation, we have

\[
\sum_{i=1}^{\ell} \hat{T}_i = \sum_{i=1}^{\ell} (\hat{V}_0^T Q_i \hat{V}_0 - w_0^T Q_i w_0 I_3)
\]

Then, with the definition of \( \hat{S} \) (3.9), we arrive at the following equivalence:

\[
\sum_{i=1}^{\ell} \hat{S}_i = \sum_{i=1}^{\ell} (Q_i - w_0^T Q_i w_0 I_4)
\]

\[
\Leftrightarrow \sum_{i=1}^{\ell} \begin{bmatrix} \hat{T}_i & 0 \\ 0 & 0 \end{bmatrix} = \sum_{i=1}^{\ell} \begin{bmatrix} (\hat{V}_0^T Q_i \hat{V} - w_0^T Q_i w_0 I_4) \\ 0 \\ 0 \end{bmatrix}
\]

\[
\Leftrightarrow \begin{bmatrix} \sum_{i=1}^{\ell} (\hat{V}_0^T Q_i \hat{V}_0 - w_0^T Q_i w_0 I_3) & 0 \\ 0 & 0 \end{bmatrix} = \sum_{i=1}^{\ell} \begin{bmatrix} (\hat{V}_0^T Q_i \hat{V} - w_0^T Q_i w_0 I_4) \\ 0 \end{bmatrix}
\]

Since \( \hat{V}_0^T w_0 = 0 \) and \( w_0 \) is an eigenvector of \( \sum_{i=1}^{\ell} Q_i \), the above equality holds.

We now prove \( \hat{S}_i \geq 0 \). By the definition of \( \hat{S} \) (3.9), it suffices to show \( \hat{T}_i \geq 0 \). Now, using the definition of \( \zeta \) (3.6) and the assumption \( \zeta \geq \ell/(\ell - 1) \), we have

\[
\hat{S}_i \geq 0 \Leftrightarrow \hat{T}_i \geq 0
\]

\[
\Leftrightarrow \frac{\zeta - \ell}{\zeta} \hat{V}_0^T Q_i \hat{V}_0 + \sum_{j=1}^{\ell} \hat{V}_0^T Q_j \hat{V}_0 - \frac{1}{\ell} \left( \sum_{j=1}^{\ell} w_0^T Q_j w_0 \right) I_3 \geq 0
\]

\[
\Leftrightarrow \frac{\sum_{j=1}^{\ell} \hat{V}_0^T Q_j \hat{V}_0}{\zeta(\ell - 1)} - \lambda_{\min} \left( \frac{1}{\ell} \sum_{j=1}^{\ell} Q_j I_3 \right) \geq 0
\]

\[
\Leftrightarrow \frac{\sum_{j=1}^{\ell} \hat{V}_0^T Q_j \hat{V}_0}{\ell} - \frac{1}{\ell} \lambda_{\min} \left( \sum_{j=1}^{\ell} Q_j I_3 \right) \geq 0,
\]
which holds true, as the minimum eigenvalue of $\sum_{j=1}^{\ell} \hat{V}_0^\top Q_j \hat{V}_0$ is exactly the second smallest eigenvalue of $\sum_{i=1}^{\ell} Q_i$ (by the definition of $\hat{V}_0$). This proves $\hat{S}_i \succeq 0$.

Finally, it remains to prove the last inequality:

$$\hat{S}_i + c_i^2 I_4 - Q_i \succ 0 \iff z_i^\top \left( \hat{S}_i + c_i^2 I_4 - Q_i \right) z_i \geq 0, \quad \forall z_i \in \mathbb{R}^4$$

(A.1)

In the above, we changed the coordinates by defining $a_i = \hat{V} z_i \in \mathbb{R}^4$. The left-hand side of (A.1) is a quadratic polynomial in entries of $a_i = [\alpha_i; a_{4i}]$, and we see in (A.1) that $a_{4i}$ appears only in the term $(c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0) a_{4i}^2 - 2a_{4i} \hat{w}_0^\top Q_i \hat{V}_0 \alpha_i$. Since $c_i^2 > \hat{w}_0^\top Q_i \hat{w}_0$, the left-hand side of (A.1) is minimized as a function of $a_{4i}$ at $a_{4i} = \hat{w}_0^\top Q_i \hat{V}_0 \alpha_i / (c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0)$. Thus, with $C_i := c_i^2 I_3 - \frac{\hat{V}_0^\top Q_i \hat{V}_0 \alpha_i}{c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0}$, we get

$$a_i^\top \left( \begin{bmatrix} \hat{T}_i & 0 \\ 0 & 0 \end{bmatrix} + c_i^2 I_4 - \hat{V}^\top Q_i \hat{V} \right) a_i \geq \alpha_i^\top \left( \hat{T}_i + C_i - \frac{\hat{V}_0^\top Q_i \hat{V}_0}{c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0} \right) \alpha_i$$

with the equality attained at $a_{4i} = \hat{w}_0^\top Q_i \hat{V}_0 \alpha_i / (c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0)$. This proves

$$\hat{S}_i + c_i^2 I_4 - Q_i \succ 0 \iff \hat{T}_i + C_i - \frac{\hat{V}_0^\top Q_i \hat{V}_0}{c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0} \succ 0$$

and we will prove the latter, which with the definition of $\hat{T}_i$ (3.10) can be written as

$$C_i - \frac{\sum_{i=1}^{\ell} \hat{w}_0^\top Q_i \hat{w}_0}{\ell} I_3 + \frac{\sum_{j \neq i} \hat{V}_0^\top Q_j \hat{V}_0 - \hat{V}_0^\top Q_i \hat{V}_0}{\zeta(\ell - 1)} \succ 0$$

Note that the above can be written as $\hat{V}_0^\top (\cdot) \hat{V}_0 \succ 0$, so a sufficient condition for the above to hold is $(\cdot) \succ 0$; in other words, it suffices to prove

$$c_i^2 I_3 - \frac{Q_i \hat{w}_0 \hat{w}_0^\top Q_i}{c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0} - \frac{\sum_{i=1}^{\ell} \hat{w}_0^\top Q_i \hat{w}_0}{\ell} I_3 - \frac{\sum_{j \neq i} (Q_j - Q_i)}{\zeta(\ell - 1)} \succ 0$$

Since $\|Q_i \hat{w}_0\|_2^2 \geq \lambda_{\max}(Q_i \hat{w}_0 \hat{w}_0^\top Q_i)$, the above holds whenever we have

$$c_i^2 - \frac{\|Q_i \hat{w}_0\|_2^2}{c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0} - \sum_{i=1}^{\ell} \frac{\|Q_i \hat{w}_0\|_2^2}{\ell} - \frac{\lambda_{\max}(\sum_{j \neq i} (Q_j - Q_i))}{\zeta(\ell - 1)} > 0$$

Using the definition of $d_i$ in (3.7), the above can be written as

$$c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0 - \frac{\|Q_i \hat{w}_0\|_2^2}{c_i^2 - \hat{w}_0^\top Q_i \hat{w}_0} - d_i > 0.$$
Multiplying the above inequality by $c_i^2 - \mathbf{w}_0^\top Q_i \mathbf{w}_0$ leads to a quadratic inequality in $c_i^2 - \mathbf{w}_0^\top Q_i \mathbf{w}_0$. Solve it for $c_i^2 - \mathbf{w}_0^\top Q_i \mathbf{w}_0$ and we get equivalently

$$c_i^2 > \mathbf{w}_0^\top Q_i \mathbf{w}_0 + \frac{d_i + \sqrt{d_i^2 + 4 \|Q_i \mathbf{w}_0\|^2}}{2},$$

which holds true, as long as our assumption (3.7) holds. This finishes the proof. \hfill \blacksquare

Appendix B. Supplementary Material: Detailed Comparison to Yang & Carlone [59].
We mentioned that our (SDR) can be treated as equivalent to the naive relaxation of Yang & Carlone (see Proposition 6 of [59]). Here we review their relaxation, highlight the subtle differences, and conclude that our relaxation is tight if and only if so is theirs.

Their derivation starts with (TLS-R) and first arrives at (an equivalent version of) (TLS-Q) via the use of quaternion products, which though could be avoided in view of relation (1.2). Then they defined $\theta_i' \in \{\pm 1\}$ and $a_i := \theta_i' w_0$ (so $\theta_i' = a_i^\top w_0$), leading to the following equivalence (as we simplify and translate):

$$(\text{TLS-Q}) \iff \min_{\mathbf{w}_0 \in \mathbb{S}^3, \theta_i' \in \{\pm 1\}} \frac{\ell}{2} \sum_{i=1}^\ell \frac{1 + \theta_i'}{2} w_0^\top Q_i w_0 + \frac{\ell}{2} \sum_{i=1}^\ell \frac{1 - \theta_i'}{2} c_i^2$$

$$\iff \min_{\mathbf{w}_0 \in \mathbb{S}^3, \theta_i' \in \{\pm 1\}} \frac{\ell}{2} \sum_{i=1}^\ell w_0^\top Q_i w_0 + \frac{\ell}{2} \sum_{i=1}^\ell a_i^\top Q_i a_i + \frac{\ell}{2} \sum_{i=1}^\ell (Q_i - c_i^2 I_4) w_0 + \frac{\ell}{2} \sum_{i=1}^\ell c_i^2$$

Define $a := [w_0; a_1; \ldots; a_\ell]$, use our definition (2.9) of $Q$, and let $Q' \in \mathbb{R}^{(\ell+1)\times(\ell+1)}$ be a block diagonal matrix with $[Q']_{00} = 0$ and $[Q']_{ii} = Q_i$. Then (B) becomes

$$(\text{QCQP-YC}) \begin{align*}
\min_{a \in \mathbb{R}^{(\ell+1)}} & \quad \frac{\ell}{2} \text{tr}(Q'aa^\top) + \frac{\ell}{2} \text{tr}(Qa^\top a) + \frac{\ell}{2} \sum_{i=1}^\ell c_i^2 \\
\text{s.t.} & \quad [aa^\top]_{00} = [aa^\top]_{ii}, \quad \forall \ i \in \{1, \ldots, \ell\} \\
& \quad \text{tr}([aa^\top]_{00}) = 1
\end{align*}$$

Here, we used the notation $Q$ defined in (2.9). To summarize, (QCQP-YC) is equivalent to the non-convex QCQP of Yang & Carlone [59] (Proposition 4 [59]). Moreover, (QCQP-YC) is also equivalent to our (QCQP), meaning that globally minimizing one of them solves the other and their minimum values $\hat{g}_{\text{QCQP}}$ and $\hat{g}_{\text{QCQP-YC}}$ are both equal to the minimum $\hat{g}_{\text{TLS-Q}}$ of (TLS-Q). Relax (QCQP-YC), and we get

$$(\text{SDR-YC}) \begin{align*}
\min_{a \geq 0} & \quad \frac{\ell}{2} \text{tr}(Q'a) + \frac{\ell}{2} \text{tr}(Qa) + \frac{\ell}{2} \sum_{i=1}^\ell c_i^2 \\
\text{s.t.} & \quad [a]_{00} = [a]_{ii}, \quad \forall \ i \in \{1, \ldots, \ell\} \\
& \quad \text{tr}([a]_{00}) = 1
\end{align*}$$

(SDR-YC) is equivalent to the naive relaxation of Yang & Carlone [59] in their Proposition 6. Not very obvious is the equivalence between (SDR) and (SDR-YC):
Proposition B.1. Assume \( c_i^2 \neq \lambda_{\min}(Q_i), \ c_i^2 \neq \lambda_{\max}(Q_i), \forall i = 1, \ldots, \ell. \) If (TLS-Q) rejects all outliers and keeps all inliers, then (SDR) is tight \( \iff \) (SDR-YC) is tight.

Proof. As before, assume that the first \( k^* \) point pairs are inliers. We first prove the \( \Rightarrow \) direction. Assume (SDR-YC) is tight, i.e., it admits \( \hat{a} \hat{a}^\top \) as an global minimizer, where \( \hat{a} \) minimizes (QCQP-YC) (Definition 2.3). Since (TLS-Q) rejects all outliers and keeps all inliers, \( \hat{a} \) is of the form \( [\hat{w}_0; \ldots; \hat{w}_0; \ldots; 0_\ell] \); here \( \hat{w}_0 \) is an eigenvector of \( \sum_{i=1}^{k^*} Q_i \) corresponding to its minimum eigenvalue \( \sum_{i=1}^{k^*} \hat{w}_0^\top Q_i \hat{w}_0 \) and \( \hat{w}_0 \) globally minimizes (TLS-Q). Moreover, we have \( \hat{y}_{\text{QCQP}} = (\sum_{i=1}^{k^*} \hat{w}_0^\top Q_i \hat{w}_0) + \sum_{j=k^*+1}^{\ell} c_j^2 \). Since (SDR-YC) is tight, the dual program of (SDR-YC), or (derivation omitted)

\[
\begin{align*}
\text{(D-YC)} & \quad \max_{\mu, \gamma} \mu + \frac{\sum_{i=1}^{\ell} c_i^2}{2} \quad \text{s.t.} \quad \frac{1}{2} Q' + \frac{1}{2} Q - \mu B - \gamma \succeq 0,
\end{align*}
\]

has the minimum value \( \sum_{i=1}^{k^*} \hat{w}_0^\top Q_i \hat{w}_0 + \sum_{j=k^*+1}^{\ell} c_j^2 \); here \( B \) was defined in Lemma 2.2 and \( \gamma \) is the \( 4(\ell + 1) \times 4(\ell + 1) \) block diagonal matrix satisfying

\[
\gamma = \begin{cases} 
\gamma'_{00} + \sum_{i=1}^{\ell} |\gamma'_{ii}| = 0, \quad \forall i \in \{1, \ldots, \ell\} \\
\text{all other entries of } \gamma \text{ are zero.}
\end{cases}
\]

In particular, the optimal \( \hat{\gamma} \) is of the form (B.1) that fulfills the optimality conditions:

\[
\begin{align*}
\left( \frac{1}{2} (Q' + Q) - \left( \sum_{i=1}^{k^*} \hat{w}_0^\top Q_i \hat{w}_0 + \sum_{j=k^*+1}^{\ell} c_j^2 \right) \right) \hat{a} &= 0 \\
\frac{1}{2} (Q' + Q) - \left( \sum_{i=1}^{k^*} \hat{w}_0^\top Q_i \hat{w}_0 + \sum_{j=k^*+1}^{\ell} c_j^2 \right) B + \hat{\gamma} &\succeq 0.
\end{align*}
\]

Similarly to (O1) of Proposition 2.5, the first condition of (B.2) is equivalent to

\[
\left( 3Q_i - c_i^2 I_4 - 4[\hat{\gamma}']_{ii} \right) \hat{w}_0 = 0, \quad i \in \{1, \ldots, k^*\}
\]

\[
\left( Q_j + c_j^2 I_4 - 4[\hat{\gamma}]_{jj} \right) \hat{w}_0 = 0, \quad j \in \{k^*+1, \ldots, \ell\}
\]

We will prove \( \hat{w} \hat{w}^\top \) globally minimizes (SDR); here \( \hat{w} := [\hat{w}_0; \ldots; \hat{w}_0; 0; \ldots; 0] \). Let \( [\hat{D}]_{00} := Q_i - 2[\hat{\gamma}']_{ii} \) for every \( i = 1, \ldots, \ell \) and \( \hat{\mu} := \sum_{i=1}^{k^*} \hat{w}_0^\top Q_i \hat{w}_0 - \sum_{i=1}^{k^*} c_i^2 \). It now suffices to show that the stationary condition (O1) and dual feasibility condition (O2) of Proposition 2.5 are fulfilled. Note that (O1) follows directly from (B.3) and the definition of \( [\hat{D}]_{00} \). It remains to prove that the second condition of (B.2) implies (O2). First observe that the second condition of (B.2), multiplied by 4, implies (\forall z_i \in \mathbb{R}^4)

\[
\begin{align*}
\left( z_0^\top \left( -4\hat{\mu} I_4 - 2 \sum_{i=1}^{\ell} c_i^2 I_4 - 4\hat{\gamma}'_{00} \right) z_0 + \right. \\
\sum_{i=1}^{\ell} z_i^{-1} \left( 2Q_i - 4[\hat{\gamma}']_{ii} \right) z_i - \\
\left. \sum_{i=1}^{\ell} 2z_0^\top (c_i^2 I_4 - Q_i) z_i \right) &\geq 0
\end{align*}
\]

Similarly to Lemma B.2, since \( c_i^2 \neq \lambda_{\min}(Q_i) \) and \( c_i^2 \neq \lambda_{\max}(Q_i) \), we have \( 2[\hat{D}]_{00} := 2Q_i - 4[\hat{\gamma}']_{ii} > 0 \). Now, substitute the minimizer \( z_i = (2[\hat{D}]_{00})^{-1} (c_i^2 I_4 - Q_i) z_0 \) into (B.4), and we see that (B.4) divided by 4 is equivalent to (\forall z_0 \in \mathbb{R}^4)

\[
\begin{align*}
\left( -\hat{\mu} I_4 - \sum_{i=1}^{\ell} c_i^2 I_4 - 2\hat{\gamma}'_{00} - \sum_{i=1}^{\ell} \frac{(c_i^2 I_4 - Q_i) (2[\hat{D}]_{00})^{-1} (c_i^2 I_4 - Q_i)}{4} \right) z_0 \geq 0,
\end{align*}
\]
or equivalently (with $z_0$ now removed)

$$-\bar{\mu}I_4 - \sum_{i=1}^\ell \frac{c_i^2 I_4}{2} - [\hat{\mathcal{V}}]_{00} - \sum_{i=1}^\ell \frac{(c_i^2 I_4 - Q_i)(2[\hat{\mathcal{V}}]_{0i})^{-1}(c_i^2 I_4 - Q_i)}{4} \preceq 0$$

$$\iff -\bar{\mu}I_4 - \sum_{i=1}^\ell \frac{(c_i^2 I_4 - Q_i)}{4} + (c_i^2 I_4 - Q_i)(2[\hat{\mathcal{V}}]_{0i})^{-1}(c_i^2 I_4 - Q_i) \preceq 0$$

$$\iff -\bar{\mu}I_4 - \sum_{i=1}^\ell \frac{2c_i^2 I_4 - 4[\hat{\mathcal{V}}]_{ii} + (c_i^2 I_4 - Q_i)(2[\hat{\mathcal{V}}]_{0i})^{-1}(c_i^2 I_4 - Q_i)}{4} \preceq 0$$

$$\iff -\bar{\mu}I_4 - \sum_{i=1}^\ell \frac{2[\hat{\mathcal{V}}]_{0i} - Q_i + c_i^2 I_4}(2[\hat{\mathcal{V}}]_{0i})^{-1}(2[\hat{\mathcal{V}}]_{0i} - Q_i + c_i^2 I_4) \preceq 0,$$

or equivalently (with $z_0$ back)

$$z_0^\top \left( -\bar{\mu}I_4 - \sum_{i=1}^\ell \frac{2[\hat{\mathcal{V}}]_{0i} - Q_i + c_i^2 I_4)(2[\hat{\mathcal{V}}]_{0i})^{-1}(2[\hat{\mathcal{V}}]_{0i} - Q_i + c_i^2 I_4)}{4} \right) z_0 \geq 0$$

or equivalently (with all other $z_i$’s back)

(B.5) $$\bar{\mu} \left\| z_0 \right\|^2 + 2 \sum_{i=1}^\ell z_i^\top [\hat{\mathcal{V}}]_{0i} z_i - \sum_{i=1}^\ell z_i^\top \left( 2[\hat{\mathcal{V}}]_{0i} - Q_i + c_i^2 I_4 \right) z_i \geq 0,$$

which is exactly (O2). The above derivation also implies that the $\Rightarrow$ direction is true. ■

**Lemma B.2.** Condition (O2) implies $[\hat{\mathcal{V}}]_{0i} \succeq 0$ for every $i = 1, \ldots, \ell$. Moreover, if $c_i^2 \neq \lambda_{\min}(Q_i)$ and $c_i^2 \neq \lambda_{\max}(Q_i)$ then (O2) implies $[\hat{\mathcal{V}}]_{0i} \succ 0$.

**Lemma B.2.** Substitute $z_0 = z_1 = \cdots = z_{\ell-1} = 0$ into (O2) and we have $z_\ell^\top [\hat{\mathcal{V}}]_{0\ell} z_\ell \geq 0$. This proves that (O2) implies $[\hat{\mathcal{V}}]_{0i} \succeq 0$ for every $i = 1, \ldots, \ell$.

Assume If $c_i^2 \neq \lambda_{\min}(Q_i)$ and $c_i^2 \neq \lambda_{\max}(Q_i)$ and $[\hat{\mathcal{V}}]_{0i}$ has an eigenvalue 0 with $[\hat{\mathcal{V}}]_{0i} z_i' = 0$ for some $z_i' \in S^3$. With some $t \in \mathbb{R}$, substitute $z_0 \in S^3$ and $z_1 = tz_1'$ and $z_2 = \cdots = z_\ell = 0$ into (O2), and we get $-\bar{\mu} + tz_0^\top (c_i^2 I_4 - Q_i) z_i' \geq 0$ for every $z_0 \in S^3$. Since $c_i^2$ is not equal to any eigenvalue of $Q_1$ (Lemma 2.1), the matrix $c_i^2 I_4 - Q_1$ is of full rank, and we have $(c_i^2 I_4 - Q_1) z_i' \neq 0$. Choose $z_0$ such that $z_0^\top (c_i^2 I_4 - Q_1) z_i' < 0$, and choose $t \to \infty$, then we see that $-\bar{\mu} + tz_0^\top (c_i^2 I_4 - Q_1) z_i' < 0$, a contradiction. This proves that (O2) and $c_i^2 \neq \lambda_{\min}(Q_i)$ and $c_i^2 \neq \lambda_{\max}(Q_i)$ imply $[\hat{\mathcal{V}}]_{0i} \succ 0$. ■

**REFERENCES**

[1] S. Agostinho, J. Gomes, and A. Del Bue, *CvxPnP: A unified convex solution to the absolute pose estimation problem from point and line correspondences*, tech. report, arXiv:1907.10545v2 [cs.CV], 2019.
C. Aholt, S. Agarwal, and R. Thomas, A QCQP approach to triangulation, in European Conference on Computer Vision, Springer, 2012, pp. 654–667.

Y. Alfassi, D. Keren, and B. Reznick, The non-tightness of a convex relaxation to rotation recovery, Sensors, 21 (2021), p. 7358.

P. Antonante, V. Tzoumas, H. Yang, and L. Carlone, Outlier-robust estimation: Hardness, minimally tuned algorithms, and applications, IEEE Transactions on Robotics, (2021).

A. S. Bandeira, N. Boumal, and A. Singer, Tightness of the maximum likelihood semidefinite relaxation for angular synchronization, Mathematical Programming, 163 (2017), pp. 145–167.

N. Boumal, An introduction to optimization on smooth manifolds. Available online, 2020, http://www.nicolasboumal.net/book.

S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

J. Briales and J. Gonzalez-Jimenez, Fast global optimality verification in 3D SLAM, in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2016, pp. 4630–4636.

J. Briales and J. Gonzalez-Jimenez, Convex global 3D registration with lagrangian duality, in IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4960–4969.

J. Briales, L. Kneip, and J. Gonzalez-Jimenez, A certifiably globally optimal solution to the non-minimal relative pose problem, in IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 145–154.

L. Brynte, V. Larsson, J. P. Iglesias, C. Olsson, and F. Kahl, On the tightness of semidefinite relaxations for rotation estimation, Journal of Mathematical Imaging and Vision, 64 (2022), pp. 57–67.

A. P. Bustos and T.-J. Chin, Guaranteed outlier removal for rotation search, in IEEE International Conference on Computer Vision, 2015, pp. 2165–2173.

E. J. Candès, T. Strohmer, and V. Voroninski, Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming, Communications on Pure and Applied Mathematics, 66 (2013), pp. 1241–1274.

L. Carlone and G. C. Calafiore, Convex relaxations for pose graph optimization with outliers, IEEE Robotics and Automation Letters, 3 (2018), pp. 1160–1167.

L. Carlone, G. C. Calafiore, C. Tommolillo, and F. Dellaert, Planar pose graph optimization: Duality, optimal solutions, and verification, IEEE Transactions on Robotics, 32 (2016), pp. 545–565.

L. Carlone, D. M. Rosen, G. Calafiore, J. J. Leonard, and F. Dellaert, Lagrangian duality in 3D SLAM: Verification techniques and optimal solutions, in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2015, pp. 125–132.

K. N. Chaudhury, Y. Khoo, and A. Singer, Global registration of multiple point clouds using semidefinite programming, SIAM Journal on Optimization, 25 (2015), pp. 468–501.

Y. Chen, Y. Chi, J. Fan, C. Ma, et al., Spectral methods for data science: A statistical perspective, Foundations and Trends® in Machine Learning, 14 (2021), pp. 566–806.

Y. Cheng, J. A. Lopez, O. Camps, and M. Sznaier, A convex optimization approach to robust fundamental matrix estimation, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2170–2178.

D. Cifuentes, A convex relaxation to compute the nearest structured rank deficient matrix, SIAM Journal on Matrix Analysis and Applications, 42 (2021), pp. 708–729.

D. Cifuentes, S. Agarwal, P. A. Parrilo, and R. R. Thomas, On the local stability of semidefinite relaxations, Mathematical Programming, (2021), pp. 1–35.

T. Ding, Y. Yang, Z. Zhu, D. P. Robinson, R. Vidal, L. Kneip, and M. C. Tsakiris, Robust homography estimation via dual principal component pursuit, in IEEE Conference on Computer Vision and Pattern Recognition, 2020, pp. 6080–6089.

E. Elhamifar and R. Vidal, Sparse subspace clustering: Algorithm, theory, and applications, IEEE transactions on pattern analysis and machine intelligence, 35 (2013), pp. 2765–2781.

A. Eriksson, C. Olsson, F. Kahl, and T.-J. Chin, Rotation averaging and strong duality, in Conference on Computer Vision and Pattern Recognition, 2018, pp. 127–135.

J. Fredriksson and C. Olsson, Simultaneous multiple rotation averaging using lagrangian duality, in Asian Conference on Computer Vision, Springer, 2012, pp. 245–258.

M. Garcia-Salgueiro, J. Briales, and J. Gonzalez-Jimenez, Certifiable relative pose estimation, Image and Vision Computing, 109 (2021), p. 104142.
[27] M. Garcia-Salgueiro and J. Gonzalez-Jimenez, *Fast and robust certifiable estimation of the relative pose between two calibrated cameras*, Journal of Mathematical Imaging and Vision, 63 (2021), pp. 1036–1056.

[28] M. Giamou, Z. Ma, V. Peretroukhin, and J. Kelly, *Certifiably globally optimal extrinsic calibration from per-sensor egomotion*, IEEE Robotics and Automation Letters, 4 (2019), pp. 367–374.

[29] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2004.

[30] B. K. Horn, *Closed-form solution of absolute orientation using unit quaternions*, Journal of the Optical Society of America A, 4 (1987), pp. 629–642.

[31] J. P. Iglesias, C. Olsson, and F. Kahl, *Global optimality for point set registration using semidefinite programming*, in IEEE Conference on Computer Vision and Pattern Recognition, 2020.

[32] D. Ikami, T. Yamasaki, and K. Aizawa, *Fast and robust estimation for unit-norm constrained linear fitting problems*, in IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8147–8155.

[33] A. K. Jain, M. N. Murty, and P. J. Flynn, *Data clustering: A review*, ACM Computing Surveys, 31 (1999), pp. 264–323.

[34] F. Kahl and D. Henrion, * Globally optimal estimates for geometric reconstruction problems*, International Journal of Computer Vision, 74 (2007), pp. 3–15.

[35] Y. Khoo and A. Kapoor, *Non-iterative rigid 2D/3D point-set registration using semidefinite programming*, IEEE Transactions on Image Processing, 25 (2016), pp. 2956–2970.

[36] P.-Y. Lajoie, S. Hu, G. Beltrame, and L. Carlone, *Modeling perceptual aliasing in SLAM via discrete–continuous graphical models*, IEEE Robotics and Automation Letters, 4 (2019), pp. 1232–1239.

[37] J. B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM Journal on Optimization, 11 (2001), pp. 796–817.

[38] M. Li, G. Liang, H. Luo, H. Qian, and T. L. Lam, *Robot-to-robot relative pose estimation based on semidefinite relaxation optimization*, in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2020, pp. 4491–4498.

[39] S. Ling, *Near-optimal bounds for generalized orthogonal procrustes problem via generalized power method*, arXiv:2112.13725 [cs.IT], (2021).

[40] C. Lu, Y.-F. Liu, W.-Q. Zhang, and S. Zhang, *Tightness of a new and enhanced semidefinite relaxation for MIMO detection*, SIAM Journal on Optimization, 29 (2019), pp. 719–742.

[41] Y. Ma, S. Soatto, J. Košecká, and S. Sastry, *An Invitation to 3D Vision: From Images to Geometric Models*, vol. 26, Springer, 2004.

[42] H. Maron, N. Dym, I. Kezurer, S. Kovalsky, and Y. Lipman, *Point registration via efficient convex relaxation*, ACM Transactions on Graphics, 35 (2016).

[43] O. Ozyesil, A. Singer, and R. Basri, *Stable camera motion estimation using convex programming*, SIAM Journal on Imaging Sciences, 8 (2015), pp. 1220–1262.

[44] A. Parra Bustos and T.-J. Chin, *Guaranteed outlier removal for point cloud registration with correspondences*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 40 (2018), pp. 2868–2882.

[45] A. Parra Bustos, T.-J. Chin, A. Eriksson, H. Li, and D. Suter, *Fast rotation search with stereographic projections for 3D registration*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38 (2016), pp. 2227–2240.

[46] L. Peng, M. Fazlyab, and R. Vidal, *Semidefinite relaxations of truncated least-squares in robust rotation search: Tight or not*, in European Conference on Computer Vision, Springer, 2022, pp. 0–0.

[47] L. Peng, M. C. Tsakiris, and R. Vidal, *ARCS: Accurate rotation and correspondences search*, in IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 11153–11163.

[48] T. Probst, D. P. Paudel, A. Chhatkuli, and L. V. Gool, *Convex relaxations for consensus and non-minimal problems in 3D vision*, in IEEE/CVF International Conference on Computer Vision, 2019, pp. 10233–10242.

[49] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard, *SE-Sync: A certifiably correct algorithm for synchronization over the special euclidean group*, The International Journal of Robotics Research, 38 (2019), pp. 95–125.
[50] J. Shi, Y. Heng, and L. Carlone, *Optimal pose and shape estimation for category-level 3d object perception*, in Robotics: Science and Systems, 2021.

[51] N. Z. Shor, *Dual quadratic estimates in polynomial and boolean programming*, Annals of Operations Research, 25 (1990), pp. 163–168.

[52] C. Tat-Jun, C. Zhipeng, and F. Neumann, *Robust fitting in computer vision: Easy or hard?*, International Journal of Computer Vision, 128 (2020), pp. 575–587.

[53] Y. Tian, K. Khosoussi, D. M. Rosen, and J. P. How, *Distributed certifiably correct pose-graph optimization*, IEEE Transactions on Robotics, 37 (2021), pp. 2137–2156.

[54] J. Unnikrishnan, S. Haghighatshoar, and M. Vetterli, *Unlabeled sensing with random linear measurements*, IEEE Transactions on Information Theory, 64 (2018), pp. 3237–3253.

[55] M. J. Wainwright, *High-Dimensional Statistics: A Non-Asymptotic Viewpoint*, Cambridge University Press, 2019.

[56] L. Wang and A. Singer, *Exact and stable recovery of rotations for robust synchronization*, Information and Inference: A Journal of the IMA, 2 (2013), pp. 145–193.

[57] E. Wise, M. Giamou, S. Khoubiyarian, A. Grover, and J. Kelly, *Certifiably optimal monocular hand-eye calibration*, in IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, IEEE, 2020, pp. 271–278.

[58] H. Yang, P. Antonante, V. Tzoumas, and L. Carlone, *Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection*, IEEE Robotics and Automation Letters, 5 (2020), pp. 1127–1134.

[59] H. Yang and L. Carlone, *A quaternion-based certifiably optimal solution to the Wahba problem with outliers*, in IEEE International Conference on Computer Vision, 2019, pp. 1665–1674.

[60] H. Yang and L. Carlone, *Certifiable outlier-robust geometric perception: Exact semidefinite relaxations and scalable global optimization*, tech. report, arXiv:2109.03349 [cs.CV], 2021.

[61] H. Yang, J. Shi, and L. Carlone, *TEASER: Fast and certifiable point cloud registration*, IEEE Transactions on Robotics, 37 (2021), pp. 314–333.

[62] J. Zhao, *An efficient solution to non-minimal case essential matrix estimation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, (2020).

[63] J. Zhao, W. Xu, and L. Kneip, *A certifiably globally optimal solution to generalized essential matrix estimation*, in IEEE Conference on Computer Vision and Pattern Recognition, 2020.

[64] Y. Zhong and N. Boumal, *Near-optimal bounds for phase synchronization*, SIAM Journal on Optimization, 28 (2018), pp. 989–1016.