Semidefinite Relaxations of Truncated Least-Squares in Robust Rotation Search: Tight or Not

Liangzu Peng[1,2], Mahyar Fazlyab[1,2], and René Vidal[2,3]

[1] Department of Electrical and Computer Engineering, Johns Hopkins University
[2] Mathematical Institute for Data Science, Johns Hopkins University
[3] Department of Biomedical Engineering, Johns Hopkins University
{lpeng25,mahyarfazlyab,rvidal}@jhu.edu

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 or not 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.

1 Introduction

Robust geometric estimation problems in computer vision have been studied for decades [28,40]. However, the analysis of their computational complexity is not sufficiently well understood [51]: There are fast algorithms that run in real time [21,43,56,46], and there are computational complexity theorems that negate the existence of efficient algorithms [51,4].[4] For example, the commonly used *consensus maximization* formulation (for *robust fitting*) is shown to be NP hard in general [51], 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 [35,57,58] 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 [33,2,24,18,41,15,14,34,7,8,9,47,27,1,37,61,25,26,3,49], much fewer works [16,42,23,48,30,55,60,52,38][5] provide theoretical insights on the robustness of semidefinite relaxations to noise, a few semidefinite relaxations [13,35,57,58] are empirically robust to outliers, and only one paper on *rotation synchronization* [54] gives theoretical guarantees for noise, outliers, and both. Complementary to the story of [51] and

[4] The catch is that the fast methods might not always be correct (e.g., at extreme outlier rates).
[5] [12,5,62,39] analyzed SDRs under noise but they are not for geometric vision problems.

inheriting the spirit of [54], in this paper we consider the question of *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 [44,43,57] for what has motivated this problem):

Problem 1 (Robust Rotation Search). Let $\{(\boldsymbol{y}_i, \boldsymbol{x}_i)\}_{i=1}^{\ell}$ be a collection of ℓ 3D point pairs. Assume that a subset $\mathcal{I}^* \subseteq \{1, \dots, \ell\}$ of these pairs are related by a 3D rotation $\boldsymbol{R}_0^* \in \mathrm{SO}(3)$ up to bounded noise $\{\boldsymbol{\epsilon}_i : \|\boldsymbol{\epsilon}_i\|_2 \leq \delta\}_{i=1}^{\ell} \subset \mathbb{R}^3$ with $\delta \geq 0$, i.e.,

$$\begin{cases} \boldsymbol{y}_i = \boldsymbol{R}_0^* \boldsymbol{x}_i + \boldsymbol{\epsilon}_i, & i \in \mathcal{I}^* \\ \boldsymbol{y}_i \text{ and } \boldsymbol{x}_i \text{ are arbitrary} & i \notin \mathcal{I}^* . \end{cases} \tag{1}$$

Here, $\mathcal{I}^*$ is called the *inlier* index set. If $i \in \mathcal{I}^*$ then $\boldsymbol{x}_i$, $\boldsymbol{y}_i$, or $(\boldsymbol{y}_i, \boldsymbol{x}_i)$ is called an inlier, otherwise it is called an *outlier*. The goal is to find $\boldsymbol{R}_0^*$ and $\mathcal{I}^*$ from $\{(\boldsymbol{y}_i, \boldsymbol{x}_i)\}_{i=1}^{\ell}$.

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*:

$$\min_{\boldsymbol{R}_0 \in \mathrm{SO}(3)} \sum_{i=1}^{\ell} \min \left\{ \left\| \boldsymbol{y}_i - \boldsymbol{R}_0 \boldsymbol{x}_i \right\|_2^2, \ c_i^2 \right\}. \tag{TLS-R}$$

While (TLS-R) is highly robust to outliers and noise [58], it is non-convex and hard to solve. Via a remarkable sequence of algebraic manipulations, [57] showed that (TLS-R) is equivalent to some non-convex *quadratically constrained quadratic program* (QCQP), which can be relaxed to 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 1). Here, we make the following contributions:

- For noiseless point sets without outliers ($\boldsymbol{\epsilon}_i = 0, \mathcal{I}^* = \{1, \dots, \ell\}$ in Problem 1), we prove that (SDR) is always tight (Theorem 1). While this result had already been proven in [57, Section E.3], our proof is simpler and shorter.
- For noiseless point sets with outliers, Theorem 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 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 $\boldsymbol{R}_0^*$. Different from Theorem 1, outliers and improper choices of c_i^2 might actually undermine the tightness of (SDR).
- For noisy point sets without outliers, Theorems 4 and 5 show that (SDR) is tight for sufficiently small noise and for sufficiently large c_i^2. Theorem 4 is not hard to prove within our analysis framework, while Theorem 5 improves over Theorem 4 by giving a better bound on c_i^2 through non-trivial constructive arguments.
- 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 6). Thus, we will discuss this case only sparingly.

Paper Organization. In Section 2 we review the derivations of (SDR) from (TLS-R) [57], while we also provide new insights. In Section 3, we discuss our main results. In Section 4, we present limitations of our work and potential avenues for future research. The proofs of our results can be found in our full paper [45].

Notations and Basics. We employ the MATLAB notation $[a_1; \dots; a_\ell]$ to denote concatenation into a column vector. Given a $4(\ell+1) \times 4(\ell+1)$ matrix $\mathcal{A}$, we employ the bracket notation $[\mathcal{A}]_{ij}$ of [57] to denote the 4×4 submatrix of $\mathcal{A}$ whose rows are indexed by $\{4i+1, \dots, 4i+4\}$ and columns by $\{4j+1, \dots, 4j+4\}$. Following our previous work on robust rotation search [46], we treat *unit quaternions* as unit vectors on the 3-sphere $\mathbb{S}^3$. Each $\boldsymbol{R} \in \mathrm{SO}(3)$ can be equivalently written as

$$\boldsymbol{R} = \begin{bmatrix} w_1^2 + w_2^2 - w_3^2 - w_4^2 & 2(w_2w_3 - w_1w_4) & 2(w_2w_4 + w_1w_3) \\ 2(w_2w_3 + w_1w_4) & w_1^2 + w_3^2 - w_2^2 - w_4^2 & 2(w_3w_4 - w_1w_2) \\ 2(w_2w_4 - w_1w_3) & 2(w_3w_4 + w_1w_2) & w_1^2 + w_4^2 - w_2^2 - w_3^2 \end{bmatrix}, \quad (2)$$

where $\boldsymbol{w} = [w_1; w_2; w_3; w_4] \in \mathbb{S}^3$ and $-\boldsymbol{w}$ are unit quaternions. Conversely, every 3×3 matrix of the form (2) with $[w_1; w_2; w_3; w_4] \in \mathbb{S}^3$ is a 3D rotation. This means a two-to-one correspondence between unit quaternions ($\mathbb{S}^3$) and 3D rotations ($\mathrm{SO}(3)$).

2 (TLS-R) and Its Relaxation: Review and New Insights

In Section 2.1 we derive a semidefinite relaxation (SDR) from (TLS-R). More specifically, we show that (TLS-R) is equivalent to a truncated least-squares problem named (TLS-Q), with the optimization variable being a unit quaternion. We further show that (TLS-Q) can be equivalently written as a quadratically constrained quadratic program, labeled as (QCQP). Thus, we obtain the semidefinite relaxation (SDR) of (QCQP), as a result of *lifting*, and their dual program (D). See Table 1 for an overview.

While Section 2.1 follows the development of [57] in spirit, our derivation is simpler. For example, we dispensed with the use of *quaternion product* [29] in [57], which is a sophisticated algebraic operation. That said, it is safe to treat our (SDR) as equivalent to the *naive relaxation* of [57]; see the full paper [45] for a detailed discussion.

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

Table 1: Descriptions of different programs. TLS means Truncated Least-Squares.

Programs	Description
(TLS-R)	The TLS problem with the optimization variable being a 3D rotation
(TLS-Q)	TLS with the optimization variable being a unit quaternion, equivalent to (TLS-R)
(QCQP)	A quadratically constrained quadratic program, equivalent to (TLS-Q)
(SDR)	The semidefinite relaxation of (QCQP), obtained via *lifting*
(D)	The dual program of (SDR) and (QCQP)

2.1 Derivation

Notice that the term $\|\boldsymbol{y}_i - \boldsymbol{R}_0\boldsymbol{x}_i\|_2^2 = \|\boldsymbol{y}_i\|_2^2 + \|\boldsymbol{x}_i\|_2^2 - 2\boldsymbol{y}_i^\top \boldsymbol{R}_0\boldsymbol{x}_i$ of (TLS-R) depends linearly on the rotation $\boldsymbol{R}_0$. Moreover, each entry of a 3D rotation $\boldsymbol{R}_0$ depends quadratically on its *unit quaternion* representation $\boldsymbol{w}_0 \in \mathbb{S}^3$; recall (2). One then naturally asks whether $\|\boldsymbol{y}_i - \boldsymbol{R}_0\boldsymbol{x}_i\|_2^2$ is a quadratic form in $\boldsymbol{w}_0$; the answer is affirmative:

Lemma 1 (Rotations and Unit Quaternions). *Let $\boldsymbol{R}_0$ be a 3D rotation, then we have*

$$\left\|\boldsymbol{y}_i - \boldsymbol{R}_0\boldsymbol{x}_i\right\|_2^2 = \boldsymbol{w}_0^\top \boldsymbol{Q}_i \boldsymbol{w}_0, \tag{3}$$

where $\boldsymbol{Q}_i$ is a 4×4 positive semidefinite matrix, and $\boldsymbol{w}_0 \in \mathbb{S}^3$ is the unit quaternion representation of $\boldsymbol{R}_0$. Moreover, the eigenvalues of $\boldsymbol{Q}_i$ are respectively

$$\left(\left\|\boldsymbol{y}_i\right\|_2 + \left\|\boldsymbol{x}_i\right\|_2\right)^2, \left(\left\|\boldsymbol{y}_i\right\|_2 + \left\|\boldsymbol{x}_i\right\|_2\right)^2, \left(\left\|\boldsymbol{y}_i\right\|_2 - \left\|\boldsymbol{x}_i\right\|_2\right)^2, \left(\left\|\boldsymbol{y}_i\right\|_2 - \left\|\boldsymbol{x}_i\right\|_2\right)^2. \tag{4}$$

While the exact form of $\boldsymbol{Q}_i$ is complicated, Lemma 1 provides a characterization of the eigenvalues of $\boldsymbol{Q}_i$, which is much easier to work with. Note that, while the relationship between 3D rotations and unit quaternions is well known (see, e.g., [29]), we have not found (4) in the literature, except in the appendix of our prior work [46].

From Lemma 1, we now see that (TLS-R) is equivalent to

$$\min_{\boldsymbol{w}_0 \in \mathbb{S}^3} \sum_{i=1}^{\ell} \min\left\{\boldsymbol{w}_0^\top \boldsymbol{Q}_i \boldsymbol{w}_0,\ c_i^2\right\}. \tag{TLS-Q}$$

Using the following simple equality ($\theta \in \{-1, 1\}$ in [35,57]; see also [31])

$$\min\{a, b\} = \min_{\theta \in \{0,1\}} \theta a + (1-\theta) b = \min_{\theta^2 = \theta} \theta a + (1-\theta) b, \tag{5}$$

problem (TLS-Q) can be equivalently written as

$$\min_{\boldsymbol{w}_0 \in \mathbb{S}^3, \theta_i^2 = \theta_i} \sum_{i=1}^{\ell} \left(\theta_i \boldsymbol{w}_0^\top \boldsymbol{Q}_i \boldsymbol{w}_0 - \theta_i c_i^2\right) + \sum_{i=1}^{\ell} c_i^2. \tag{6}$$

Note that, while the constant $\sum_{i=1}^{\ell} c_i^2$ in (6) can be ignored, keeping it there will simplify matters. Even though the objective of (6) is a cubic polynomial in the entries of the unknowns $\boldsymbol{w}_0$ and θ_i's, problem (6) is equivalent to a quadratic program. Indeed, let $\boldsymbol{w}_i := \theta_i \boldsymbol{w}_0$, which implies $\theta_i = \boldsymbol{w}_0^\top \boldsymbol{w}_i$. Then (6) becomes

$$\min_{\boldsymbol{w}_0 \in \mathbb{S}^3, \boldsymbol{w}_i \in \{\boldsymbol{w}_0, 0\}} \sum_{i=1}^{\ell} \left(\boldsymbol{w}_0^\top (\boldsymbol{Q}_i - c_i^2 \boldsymbol{I}_4) \boldsymbol{w}_i\right) + \sum_{i=1}^{\ell} c_i^2 \tag{7}$$

The objective function of problem (7) is now quadratic. Moreover, the constraints are also quadratic. To see this, one easily verifies that the binary constraint $\boldsymbol{w}_i \in \{\boldsymbol{w}_0, 0\}$

can be equivalently written quadratically as $\boldsymbol{w}_i \boldsymbol{w}_0^\top = \boldsymbol{w}_i \boldsymbol{w}_i^\top$. Collecting all vectors of variables into a $4(\ell+1)$ dimensional column vector $\mathcal{w} = [\boldsymbol{w}_0; \dots; \boldsymbol{w}_\ell]$, we have

$$\begin{cases} \boldsymbol{w}_0 \in \mathbb{S}^3 \\ \boldsymbol{w}_i \in \{\boldsymbol{w}_0, 0\} \end{cases} \Leftrightarrow \begin{cases} \operatorname{tr}(\boldsymbol{w}_0 \boldsymbol{w}_0^\top) = 1 \\ \boldsymbol{w}_i \boldsymbol{w}_0^\top = \boldsymbol{w}_i \boldsymbol{w}_i^\top \end{cases} \Leftrightarrow \begin{cases} \operatorname{tr}\left([\mathcal{w}\mathcal{w}^\top]_{00}\right) = 1 \\ [\mathcal{w}\mathcal{w}^\top]_{0i} = [\mathcal{w}\mathcal{w}^\top]_{ii}. \end{cases} \tag{8}$$

In the last equivalence of (8) we used the notation $[\,\cdot\,]_{ij}$ of Section 1. Having confirmed (8), we can now equivalently transform (7) into the following (QCQP):

$$\min_{\mathcal{w} \in \mathbb{R}^{4(\ell+1)}} \ \operatorname{tr}\left(\mathcal{Q}\mathcal{w}\mathcal{w}^\top\right) + \sum_{i=1}^{\ell} c_i^2 \tag{QCQP}$$

$$\text{s.t. } [\mathcal{w}\mathcal{w}^\top]_{0i} = [\mathcal{w}\mathcal{w}^\top]_{ii}, \ \ \forall\, i \in \{1, \dots, \ell\} \tag{9}$$

$$\operatorname{tr}\left([\mathcal{w}\mathcal{w}^\top]_{00}\right) = 1 \tag{10}$$

In (QCQP), $\mathcal{Q}$ is our $4(\ell+1) \times 4(\ell+1)$ data matrix, symmetric and satisfying

$$\begin{cases} [\mathcal{Q}]_{0i} = [\mathcal{Q}]_{i0} = \frac{1}{2}(\boldsymbol{Q}_i - c_i^2 \boldsymbol{I}_4), \ \ \forall\, i \in \{1, \dots, \ell\} \\ \text{all other entries of } \mathcal{Q} \text{ are zero.} \end{cases} \tag{11}$$

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

$$\max_{\mu, \mathcal{D}} \ \mu + \sum_{i=1}^{\ell} c_i^2 \quad \textit{s.t.} \quad \mathcal{Q} - \mu\mathcal{B} - \mathcal{D} \succeq 0 \tag{D}$$

$$\min_{\mathcal{W} \succeq 0} \ \operatorname{tr}\left(\mathcal{Q}\mathcal{W}\right) + \sum_{i=1}^{\ell} c_i^2 \tag{SDR}$$

$$\textit{s.t. } [\mathcal{W}]_{0i} = [\mathcal{W}]_{ii}, \ \ \forall\, i \in \{1, \dots, \ell\} \tag{12}$$

$$\operatorname{tr}\left([\mathcal{W}]_{00}\right) = 1 \tag{13}$$

In the dual (D)*, $\mathcal{B}$ is a matrix of zeros except $[\mathcal{B}]_{00} = \boldsymbol{I}_4$, while $\mathcal{D} \in \mathbb{R}^{4(\ell+1)\times 4(\ell+1)}$ is a matrix of dual variables accounting for the ℓ constraints of* (9)*, i.e., $\mathcal{D}$ satisfies*

$$\begin{cases} \mathcal{D} \textit{ is symmetric}, \ \ [\mathcal{D}]_{ii} + 2[\mathcal{D}]_{0i} = 0, \ \ \forall\, i \in \{1, \dots, \ell\} \\ \textit{all other entries of } \mathcal{D} \textit{ are zero.} \end{cases} \tag{14}$$

In (D), we omitted constraint (14) on $\mathcal{D}$ for simplicity, but we will keep it in mind.

2.2 The Tightness and KKT Optimality Conditions

Studying the interplay among (D), (SDR), and (QCQP) is the main theme of the paper, which will be discussed in more detail in Section 3. Here we give some basic results, to begin with. Note that weak duality between (QCQP) and (D) holds as a result of Lagrangian calculation. Also, with $\mathcal{D}$ of the form (14) satisfying $[\mathcal{D}]_{0i} := \frac{1}{2}(\boldsymbol{Q}_i + c_i^2 \boldsymbol{I}_4)$ and with μ sufficiently small, it is not hard to show that $\mathcal{Q} - \mu\mathcal{B} - \mathcal{D} \succ 0$ (cf. the proof of Theorem 1), thus $(\mu, \mathcal{D})$ is a strictly feasible point of (D) and the Slater's condition is satisfied, hence strong duality between (D) and (SDR) holds. In summary, we have

$$\hat{\mu} + \sum_{i=1}^{\ell} c_i^2 = \hat{g}_{\mathrm{D}} = \hat{g}_{\mathrm{SDR}} \leq \hat{g}_{\mathrm{QCQP}}, \tag{15}$$

where the last three terms of (15) are the optimal objective values of (D), (SDR), and (QCQP), respectively. The next question is whether there are conditions under which the last inequality becomes an equality, i.e., $\hat{g}_{\mathrm{SDR}} = \hat{g}_{\mathrm{QCQP}}$. A (seemingly) stronger version of this objective value equality is the following notion of tightness:

Definition 1 (Tightness). *(SDR) is said to be tight if it admits $\hat{w}(\hat{w})^\top$ as a global minimizer, where $\hat{w} \in \mathbb{R}^{4(\ell+1)}$ globally minimizes (QCQP).*

The following proposition provides a starting point for tightness analysis whose proof follows from a standard duality argument:

Proposition 1 (Optimality Conditions). *Recall that $\mathcal{B}$ is defined in Lemma 2 and $\hat{w}$ denotes a global minimizer of (QCQP). Let $\hat{\mu}$ be such that $\hat{\mu} + \sum_{i=1}^{\ell} c_i^2$ globally minimizes (D). (SDR) is tight if and only if there is a matrix $\hat{\mathcal{D}}$ of the form (14) that satisfies*

(i) $(\mathcal{Q} - \hat{\mu}\mathcal{B} - \hat{\mathcal{D}})\hat{w} = 0$
(ii) $\mathcal{Q} - \hat{\mu}\mathcal{B} - \hat{\mathcal{D}} \succeq 0$
(iii) The minimum $\hat{\mu} + \sum_{i=1}^{\ell} c_i^2$ of (SDR) is also the minimum of (QCQP).

Next, we simplify the above optimality conditions to ease the use.

Proposition 2 (Simplified Optimality Conditions). *Let $\boldsymbol{Q}_1, \dots, \boldsymbol{Q}_{k^*}$ be inliers and $\boldsymbol{Q}_{k^*+1}, \dots, \boldsymbol{Q}_\ell$ outliers. Let $\hat{\boldsymbol{w}}_0$ globally minimize (TLS-Q). Assume (TLS-Q) preserves all inliers and rejects all outliers. Then the $4(\ell+1)$ dimensional vector $\hat{w} = [\hat{\boldsymbol{w}}_0; \dots; \hat{\boldsymbol{w}}_0; 0; \dots, 0]$, where $\hat{\boldsymbol{w}}_0$ appeared $k^* + 1$ times, is a global minimizer of (QCQP), and the optimality conditions of Proposition 1 can be simplified as follows:*

– *If $\hat{\mu} = \sum_{i=1}^{k^*} \left(\hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 - c_i^2\right)$, condition (i) of Proposition 1 is equivalent to*

$$\begin{cases} \left(2[\hat{\mathcal{D}}]_{0i} + \boldsymbol{Q}_i - c_i^2 \boldsymbol{I}_4\right)\hat{\boldsymbol{w}}_0 = 0, & \forall i \in \{1, \dots, k^*\} \\ \left(2[\hat{\mathcal{D}}]_{0j} + c_j^2 \boldsymbol{I}_4 - \boldsymbol{Q}_j\right)\hat{\boldsymbol{w}}_0 = 0, & \forall j \in \{k^* + 1, \dots, \ell\}. \end{cases} \tag{O1}$$

– *Condition (ii) of Proposition 1 is equivalent to ($\forall \boldsymbol{z}_i \in \mathbb{R}^4$)*

$$-\hat{\mu}\left\|\boldsymbol{z}_0\right\|_2^2 + 2\sum_{i=1}^{\ell} \boldsymbol{z}_i^\top [\hat{\mathcal{D}}]_{0i} \boldsymbol{z}_i - \sum_{i=1}^{\ell} \boldsymbol{z}_0^\top \left(2[\hat{\mathcal{D}}]_{0i} - \boldsymbol{Q}_i + c_i^2 \boldsymbol{I}_4\right)\boldsymbol{z}_i \geq 0. \tag{O2}$$

– *Condition (iii) of Proposition 1 is equivalent to*

$$\hat{\mu} = \sum_{i=1}^{k^*} \left(\hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 - c_i^2\right). \tag{O3}$$

Thanks to Propositions 1 and 2, establishing whether (SDR) is tight or not reduces to finding *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 Theorems 1-6, which we discuss in greater detail in the next section.

3 Main Results

In this section we present our main results regarding the tightness of (SDR). Our results are naturally categorized into four subsections. Section 3.1 treats the simplest noiseless + outlier-free case (Theorem 1). Sections 3.2 and 3.3 consider the case where the data is corrupted by outliers (Theorems 2 and 3) and noise (Theorems 4 and 5), respectively, and Section 3.4 brings them together for the noisy + outliers case (Theorem 6).

3.1 The Noiseless + Outlier-Free Case

Theorem 1 (Noiseless and Outlier-Free Point Sets). *In the absence of noise and outliers,* (SDR) *is tight, meaning that* $w^*(w^*)^\top$ *globally minimizes* (SDR)*, where* $w^* = [\boldsymbol{w}_0^*; \dots; \boldsymbol{w}_0^*] \in \mathbb{R}^{4(\ell+1)}$ *is a global minimizer of* (QCQP)*.*

Proof. Note that $w^* = [\boldsymbol{w}_0^*; \dots; \boldsymbol{w}_0^*]$ is a global minimizer of (QCQP) that results in the optimal value 0. Let $\hat{\mathcal{D}}$ satisfy the constraint (14) with $[\hat{\mathcal{D}}]_{0i} := \frac{1}{2}(\boldsymbol{Q}_i + c_i^2 \boldsymbol{I}_4)$ for every $i = 1, \dots, \ell$ and let $\hat{\mu} := -\sum_{i=1}^{\ell} c_i^2$. Then, with $\boldsymbol{Q}_i \boldsymbol{w}_0^* = 0$ (Lemma 1), one easily verifies that optimality conditions (O1) and (O3) of Proposition 2 hold. It remains to prove condition (O2). Substitute the values of $[\hat{\mathcal{D}}]_{0i}$, $\hat{\mu}$ into (O2) and it simplifies:

$$\sum_{i=1}^{\ell} c_i^2 \|\boldsymbol{z}_0\|_2^2 + \sum_{i=1}^{\ell} \boldsymbol{z}_i^\top (\boldsymbol{Q}_i + c_i^2 \boldsymbol{I}_4) \boldsymbol{z}_i - 2 \sum_{i=1}^{\ell} c_i^2 \boldsymbol{z}_i^\top \boldsymbol{z}_0 \geq 0, \quad \forall \boldsymbol{z}_i \in \mathbb{R}^4 \tag{16}$$

$$\Leftrightarrow \sum_{i=1}^{\ell} \left(c_i^2 \|\boldsymbol{z}_0 - \boldsymbol{z}_i\|_2^2 + \boldsymbol{z}_i^\top \boldsymbol{Q}_i \boldsymbol{z}_i \right) \geq 0, \quad \forall \boldsymbol{z}_i \in \mathbb{R}^4 \tag{17}$$

Thus (O2) holds, as every $\boldsymbol{Q}_i$ is positive semidefinite (Lemma 1). One also observes that the equality is attained if and only if $\boldsymbol{z}_0 = \cdots = \boldsymbol{z}_\ell = \boldsymbol{w}_0^*$ or $\boldsymbol{z}_0 = \cdots = \boldsymbol{z}_\ell = 0$.

Our contribution here is a shorter proof for Theorem 1 than that in [57]. Besides Lemma 1, another key idea that shortens the proof is our construction of the dual certificate $\hat{\mathcal{D}}$ (or $[\hat{\mathcal{D}}]_{0i}$'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 1), inspect the first and second order Riemannian optimality conditions (cf. [6]), discover some properties of data (e.g., Lemma 1), (3) *repeatedly try* different choices of certificates. In what follows, due to space limitations, we not always provide full proofs of our theorems, but we always provide a sketch of the dual certificate.

3.2 The Noiseless + Outliers Case

Different from Theorem 1, in this case the tightness of (SDR) depends on both the data and truncation parameter c_i^2, as stated in the following result.

Theorem 2 (Noiseless Point Sets with Outliers). *Suppose there is no noise. Consider (TLS-Q) with outliers* $\boldsymbol{Q}_{k^*+1},\dots,\boldsymbol{Q}_\ell$ $(k^* < \ell)$. *Recall* $\boldsymbol{w}_0^*$ *denotes the unit quaternion that represents the ground-truth rotation* $\boldsymbol{R}_0^*$. *Let* $w^* := [\boldsymbol{w}_0^*;\dots;\boldsymbol{w}_0^*;0;\dots;0] \in \mathbb{R}^{4(\ell+1)}$, *where* $\boldsymbol{w}_0^*$ *appears* k^*+1 *times, and let* $\mathcal{W}^* := w^*(w^*)^\top$. *Then we have:*

- *If* $0 < c_j^2 < \lambda_{\min}(\boldsymbol{Q}_j)$, *for all* $j = k^*+1,\dots,\ell$, *then (SDR) is tight, admitting* $\mathcal{W}^*$ *as a global minimizer.*
- *If* $c_j^2 > (\boldsymbol{w}_0^*)^\top \boldsymbol{Q}_j \boldsymbol{w}_0^*$ *for some* $j \in \{k^*+1,\dots,\ell\}$, *then* $\mathcal{W}^*$ *is not a global minimizer of (SDR).*

Proof (Sketch). For the first part, note that $c_j^2 < \lambda_{\min}(\boldsymbol{Q}_j)$, $\forall j = k^*+1,\dots,\ell$, so (TLS-Q) rejects all outliers and preserves all inliers, $\boldsymbol{w}_0^*$ globally minimizes (TLS-Q), and w^* globally minimizes (QCQP) with the minimum value $\sum_{j=k^*+1}^{\ell} c_j^2$. Let $[\hat{\mathcal{D}}]_{0i} := \frac{1}{2}(\boldsymbol{Q}_i + c_i^2 \boldsymbol{I}_4)$ $(\forall i = 1,\dots,k^*)$, $[\hat{\mathcal{D}}]_{0j} := \frac{1}{2}(\boldsymbol{Q}_j - c_j^2 \boldsymbol{I}_4)$ $(\forall j = k^*+1,\dots,\ell)$, and $\hat{\mu} := -\sum_{i=1}^{k^*} c_i^2$. With $\boldsymbol{Q}_i \boldsymbol{w}_0^* = 0$, $\forall i = 1,\dots,k^*$, (Lemma 1), one easily verifies conditions (O1) and (O3) of Proposition 2 hold. Condition (O2) is the same as

$$\sum_{i=1}^{k^*} \left(c_i^2 \left\| \boldsymbol{z}_0 - \boldsymbol{z}_i \right\|_2^2 + \boldsymbol{z}_i^\top \boldsymbol{Q}_i \boldsymbol{z}_i \right) + \sum_{j=k^*+1}^{\ell} \boldsymbol{z}_j^\top (\boldsymbol{Q}_j - c_j^2 \boldsymbol{I}_4) \boldsymbol{z}_j \geq 0, \quad \forall \boldsymbol{z}_i \in \mathbb{R}^4, \quad (18)$$

which holds true because $\boldsymbol{Q}_i$' are positive semidefinite as per Lemma 1 and $\boldsymbol{Q}_j \succeq c_j^2 \boldsymbol{I}_4$. This proves the first part. For the second part, it suffices to prove that, given $c_\ell^2 > (\boldsymbol{w}_0^*)^\top \boldsymbol{Q}_\ell \boldsymbol{w}_0^*$, the three conditions of Proposition 1 (or Proposition 2) can not be simultaneously satisfied by w^* and any $\hat{\mu}$ and $\hat{\mathcal{D}}$, where $\hat{\mathcal{D}}$ is of the form (14). This is proved by constructing a specific counterexample; see our full paper [45] for details.

Remark 1 (Noiseless Point Sets with Random Outliers). If outlier $(\boldsymbol{y}_j, \boldsymbol{x}_j)$ is randomly drawn from $\mathbb{R}^3 \times \mathbb{R}^3$ according to some continuous probability distribution, then with probability 1 we have $\|\boldsymbol{y}_j\|_2 \neq \|\boldsymbol{x}_j\|_2$ (Lemma 2 of [53]), which implies $\lambda_{\min}(\boldsymbol{Q}_j) > 0$. Thus, (SDR) is always tight to such random outliers, if $c_j^2 \to 0$. Note that this discussion is theoretical and does not apply to the case where $\big| \|\boldsymbol{y}_j\|_2 - \|\boldsymbol{x}_j\|_2 \big|$ is nonzero but is below machine accuracy, as c_j^2 can not be set even smaller (the case $c_j^2 = 0$ is trivial).

If the condition $c_j^2 < \lambda_{\min}(\boldsymbol{Q}_j)$ of the first statement in Theorem 2 holds then $\boldsymbol{Q}_j$ will always be rejected by (TLS-Q) as an outlier. In fact, since $\lambda_{\min}(\boldsymbol{Q}_j)$ can be easily computed (Lemma 1), in practice one usually throws away the point pairs $(\boldsymbol{y}_j, \boldsymbol{x}_j)$'s for which $c_j^2 < \lambda_{\min}(\boldsymbol{Q}_j)$ as a means of preprocessing (cf. [11,46]), and these point pairs might not enter into the semidefinite optimization. Thus, Theorem 2 suggests that (SDR) can distinguish this type of "simple" outliers, as long as c_j^2 is properly chosen.

In the second statement of Theorem 2, if the condition $c_\ell^2 > (\boldsymbol{w}_0^*)^\top \boldsymbol{Q}_\ell \boldsymbol{w}_0^*$ holds true for outlier $\boldsymbol{Q}_\ell$, then (TLS-Q) would attempt to minimize $\boldsymbol{w}_0^\top \boldsymbol{Q}_\ell \boldsymbol{w}_0 + \sum_{i=1}^{k^*} \boldsymbol{w}_0^\top \boldsymbol{Q}_i \boldsymbol{w}_0$ over $\boldsymbol{w}_0 \in \mathbb{S}^3$ at least—an outlier showed up in the eigenvalue optimization—thus the global minimizer of (TLS-Q) is unlikely to be $\boldsymbol{w}_0^*$. This is why we do not expect $\mathcal{W}^*$ to globally minimize (SDR); our theorem confirms this.

Admittedly, Theorem 2 leaves a gap: What if c_j^2 is sandwiched between $\lambda_{\text{min}}(\boldsymbol{Q}_j)$ and $(\boldsymbol{w}_0^*)^\top \boldsymbol{Q}_j \boldsymbol{w}_0^*$? Or what can we say about the tightness of (SDR) if

$$\lambda_{\text{min}}(\boldsymbol{Q}_j) < c_j^2 < (\boldsymbol{w}_0^*)^\top \boldsymbol{Q}_j \boldsymbol{w}_0^* \ \ ? \tag{19}$$

While our empirical observation suggests that $\mathcal{W}^*$ does not globally minimize (SDR) if (19) holds (with $k^* < \ell$), the analysis of this case without further assumptions on the outliers appears hard. The difficulty is that the outliers $\boldsymbol{Q}_j$'s could be so adversarial that $(\boldsymbol{w}_0^*)^\top \boldsymbol{Q}_j \boldsymbol{w}_0^*$ is arbitrarily close[6] to 0 while $\lambda_{\text{min}}(\boldsymbol{Q}_j) = 0$ for every $j > k^*$. Thus, the value of Theorem 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 $\boldsymbol{Q}_j$'s can not be simply removed by preprocessing, e.g., $\lambda_{\text{min}}(\boldsymbol{Q}_j) = 0$. In particular, we assume the outliers are *clustered* and show that the (SDR) under investigation is even more vulnerable:

Theorem 3 (Noiseless Point Sets with Clustered Outliers). *With the notation of Theorem 2, further suppose outliers $\boldsymbol{Q}_{k^*+1}, \dots, \boldsymbol{Q}_\ell$ are "clustered" in the sense that*

$$\boldsymbol{Q}_{k^*+1} \boldsymbol{w}_0^{\text{cl}} = \cdots = \boldsymbol{Q}_\ell \boldsymbol{w}_0^{\text{cl}} = 0 \tag{20}$$

with $\boldsymbol{w}_0^{\text{cl}} \in \mathbb{S}^3$ some unit quaternion that is different from $\pm \boldsymbol{w}_0^*$. If

$$1 - \frac{\sum_{j=k^*+1}^{\ell} c_j^2}{2 \sum_{i=1}^{k^*} c_i^2} < \left| (\boldsymbol{w}_0^{\text{cl}})^\top \boldsymbol{w}_0^* \right|, \tag{21}$$

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

Proof (Sketch). The proof uses the same idea as in proving the second statement of Theorem 2: Prove via counterexamples that the three conditions of Proposition 2 can no hold simultaneously. They differ though, in how the counterexamples are constructed.

The *clustered* outliers of Theorem 3 defined in the sense of (20) mean that the outlier pairs $(\boldsymbol{y}_j, \boldsymbol{x}_j)$ $(j > k^*)$ are related by the same 3D rotation $\boldsymbol{R}_0^{\text{cl}}$ that correspond to $\boldsymbol{w}_0^{\text{cl}}$, that is $\boldsymbol{y}_j = \boldsymbol{R}_0^{\text{cl}} \boldsymbol{x}_j$, $\forall j > k^*$ (Lemma 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 worse case; it should be distinguished from data clustering [32,22].

To understand condition (21) of Theorem 3, consider a situation where all truncation parameters are equal, $c_1^2 = \cdots = c_\ell^2$. Then (21) simplifies to $1 - (\ell - k^*)/(2k^*) < \left|(\boldsymbol{w}_0^{\text{cl}})^\top \boldsymbol{w}_0^*\right|$; also note that $\left|(\boldsymbol{w}_0^{\text{cl}})^\top \boldsymbol{w}_0^*\right| \in [0, 1)$. Thus, if $\ell - k^* > 2k^*$, then (21) always holds, and so $\mathcal{W}^*$ never globally minimizes (SDR), which is forgivable as in this case $\boldsymbol{w}_0^*$ 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 implies that $\mathcal{W}^*$ would still fail to globally minimize (SDR) as long as $\left|(\boldsymbol{w}_0^{\text{cl}})^\top \boldsymbol{w}_0^*\right| > 3/4$, but $\boldsymbol{w}_0^*$ would *in general* globally minimize (TLS-Q) with suitable c_j^2 (cf. [59]). Then one might conclude that (SDR) is *strictly* less robust to outliers than (TLS-Q).

[6] Alternatively, if $(\boldsymbol{w}_0^*)^\top \boldsymbol{Q}_j \boldsymbol{w}_0^*$ is small, then $\boldsymbol{Q}_j$ might be treated as noisy data rather than an outlier. We consider such noisy case in Sections 3.3 (without outliers) and 3.4 (with outliers).

Finally, we note that Theorem 3 might be overly pessimistic. In fact, experiments show that (SDR) is robust to 40%-50% outliers $(\boldsymbol{y}_j, \boldsymbol{x}_j)$'s, where $\boldsymbol{y}_j$ and $\boldsymbol{x}_j$ are sampled uniformly at random from $\mathbb{S}^2$ (so $\lambda_{\text{min}}(\boldsymbol{Q}_j) = 0$ by Lemma 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 2 and 3: (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 $\mathcal{W}^*$, the latter being much more challenging though, in our opinion.

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. Since we already have Theorem 1, one might wonder whether it can be extended to the noisy + outlier-free case using some continuity argument. In fact, [20] 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 hard to verify. In fact, it is not applicable to our case directly, as in our problem the truncation parameters c_i^2 also have impacts on the tightness, and the approach of [20] 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 by decomposing $\boldsymbol{Q}_i$ of (TLS-Q) into the pure data part and noise part :

Lemma 3. *Let $\boldsymbol{R}_0 \in \mathrm{SO}(3)$. If $(\boldsymbol{y}_i, \boldsymbol{x}_i)$ is an inlier that obeys (1), then we have*

$$\left\|\boldsymbol{y}_i - \boldsymbol{R}_0\boldsymbol{x}_i\right\|_2^2 = \boldsymbol{w}_0^\top \boldsymbol{Q}_i \boldsymbol{w}_0, \quad \boldsymbol{Q}_i = \boldsymbol{P}_i + \boldsymbol{E}_i + \left\|\boldsymbol{\epsilon}_i\right\|_2^2 \boldsymbol{I}_4, \tag{22}$$

where $\boldsymbol{w}_0 \in \mathbb{S}^3$ is the unit quaternion representation of $\boldsymbol{R}_0$, and $\boldsymbol{P}_i$ and $\boldsymbol{E}_i$ are 4×4 symmetric matrices that repsectively satisfy the following properties:

- *$\boldsymbol{P}_i$ is positive semidefinite with its entries depending on $\boldsymbol{y}_i$ and $\boldsymbol{x}_i$, and it has two different eigenvalues $4\|\boldsymbol{x}_i\|_2^2$ and 0, each of multiplicity 2. The ground-truth unit quaternion $\boldsymbol{w}_0^*$ is an eigenvector of $\boldsymbol{P}_i$ corresponding to eigenvalue 0, i.e., $\boldsymbol{P}_i\boldsymbol{w}_0^* = 0$. In particular, we have $\boldsymbol{Q}_i\boldsymbol{w}_0^* = \boldsymbol{P}_i\boldsymbol{w}_0^* = 0$ in the noiseless case.*
- *$\boldsymbol{E}_i$ has entries depending on $\boldsymbol{y}_i$, $\boldsymbol{x}_i$, and noise $\boldsymbol{\epsilon}_i$, and has two different eigenvalues $2\boldsymbol{\epsilon}_i^\top \boldsymbol{R}_0^*\boldsymbol{x}_i + 2\|\boldsymbol{\epsilon}_i\|_2\|\boldsymbol{x}_i\|_2$ and $2\boldsymbol{\epsilon}_i^\top \boldsymbol{R}_0^*\boldsymbol{x}_i - 2\|\boldsymbol{\epsilon}_i\|_2\|\boldsymbol{x}_i\|_2$, each of multiplicity 2. We have $\boldsymbol{w}_0^\top \boldsymbol{E}_i \boldsymbol{w}_0 = 2\boldsymbol{\epsilon}_i^\top(\boldsymbol{R}_0^*\boldsymbol{x}_i - \boldsymbol{R}_0\boldsymbol{x}_i)$ and in particular $(\boldsymbol{w}_0^*)^\top \boldsymbol{E}_i \boldsymbol{w}_0^* = 0$.*

A Warm-Up Result We first consider a simple case where c_i^2 is sufficiently large:

Theorem 4 (Noisy and Outlier-Free Point Sets $\{(\boldsymbol{y}_i, \boldsymbol{x}_i)\}_{i=1}^{\ell}$). *Consider (TLS-Q) with noisy inliers $\{\boldsymbol{Q}_i\}_{i=1}^{\ell}$ and $\hat{\boldsymbol{w}}_0 \in \mathbb{S}^3$ its global minimizer. Let $\hat{w} := [\hat{\boldsymbol{w}}_0; \dots; \hat{\boldsymbol{w}}_0]$. If $c_i^2 > \lambda_{\text{max}}(\boldsymbol{Q}_i)$ for every $i = 1, \dots, \ell$, then (SDR) is tight as it admits $\hat{w}(\hat{w})^\top$ as* *a global minimizer, and, moreover, the angle $\hat{\tau}_0^*$ between $\hat{\boldsymbol{w}}_0$ and the ground-truth unit quaternion $\boldsymbol{w}_0^* \in \mathbb{S}^3$ grows proportionally with the magnitude of noise $\boldsymbol{\epsilon}_i$:*

$$\sin^2(\hat{\tau}_0^*) \leq \frac{4\sum_{i=1}^{\ell} \|\boldsymbol{\epsilon}_i\|_2 \|\boldsymbol{x}_i\|_2}{\lambda_{\text{min2}}\big(\sum_{i=1}^{\ell} \boldsymbol{P}_i\big)}, \quad \sin^2(\hat{\tau}_0^*) := 1 - (\hat{\boldsymbol{w}}_0^\top \boldsymbol{w}_0^*)^2 \tag{23}$$

In (23), each $\boldsymbol{P}_i$ corresponds to the "pure data" part of $\boldsymbol{Q}_i$ that satisfies $\boldsymbol{P}_i \boldsymbol{w}_0^* = 0, \boldsymbol{P}_i \succeq 0$ (Lemma 3), and $\lambda_{\text{min2}}(\cdot)$ denotes the second smallest eigenvalue of a matrix.

Proof (Sketch). Let $\hat{\mathcal{D}}$ satisfy (14) with $[\hat{\mathcal{D}}]_{0i} = \frac{1}{2}\big(c_i^2 \boldsymbol{I}_4 - \boldsymbol{Q}_i\big), \forall i = 1, \dots, \ell$, and let $\hat{\mu} = \sum_{i=1}^{\ell} \big(\hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 - c_i^2\big)$. Since $c_i^2 > \lambda_{\text{max}}(\boldsymbol{Q}_i)$, the minimum of (QCQP) is $\sum_{i=1}^{\ell} \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0$. Again, one easily verifies conditions (O1) and (O3) of Proposition 2 hold. Moreover, condition (O2) of Proposition 2 is equivalent to ($\forall \boldsymbol{z}_i \in \mathbb{R}^4$)

$$\sum_{i=1}^{\ell} \bigg((\boldsymbol{z}_i - \boldsymbol{z}_0)^\top \big(c_i^2 \boldsymbol{I}_4 - \boldsymbol{Q}_i\big)(\boldsymbol{z}_i - \boldsymbol{z}_0) + \boldsymbol{z}_0^\top \Big(\boldsymbol{Q}_i - (\hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0)\boldsymbol{I}_4\Big)\boldsymbol{z}_0 \bigg) \geq 0,$$

which holds, as $c_i^2 \boldsymbol{I}_4 - \boldsymbol{Q}_i \succeq 0$ and $\hat{\boldsymbol{w}}_0^\top \big(\sum_{i=1}^{\ell} \boldsymbol{Q}_i\big)\hat{\boldsymbol{w}}_0$ is the minimum eigenvalue of $\sum_{i=1}^{\ell} \boldsymbol{Q}_i$. This proves the tightness. For the proof of bound (23), see our full paper [45].

First we note that the error bound (23) becomes zero as $\epsilon_i \to 0$ and thus $\hat{\boldsymbol{w}}_0 = \boldsymbol{w}_0^*$, provided that $\|\boldsymbol{x}_i\|_2 / \lambda_{\text{min2}}\big(\sum_{i=1}^{\ell} \boldsymbol{P}_i\big)$ is not too large. The denominator $\lambda_{\text{min2}}\big(\sum_{i=1}^{\ell} \boldsymbol{P}_i\big)$ seems inevitable, as it usually determines the stability of solving a minimum eigenvalue problem: If $\lambda_{\text{min2}}\big(\sum_{i=1}^{\ell} \boldsymbol{P}_i\big) \to 0 = \lambda_{\text{min}}\big(\sum_{i=1}^{\ell} \boldsymbol{P}_i\big)$, then $\hat{\boldsymbol{w}}_0$ can be arbitrarily far from $\boldsymbol{w}_0^*$ even in the slightest presence of noise. Similarly, the bound can be trivial if $\|\boldsymbol{x}_i\|_2$ is too large. However, one can show that, for ℓ large enough, if the entries of each $\boldsymbol{x}_i$ are i.i.d. Gaussian, then $\lambda_{\text{min2}}\big(\sum_{i=1}^{\ell} \boldsymbol{Q}_i\big)$ is a positive multiple of $\sum_{i=1}^{\ell} \|\boldsymbol{x}_i\|_2^2$ with high probability; see our paper [45] for rigorous statements. In other words, for random Gaussian data, $\|\boldsymbol{x}_i\|_2 / \lambda_{\text{min2}}\big(\sum_{i=1}^{\ell} \boldsymbol{P}_i\big)$ is small and the bound (23) is well-behaved.

Condition $c_i^2 > \lambda_{\text{max}}(\boldsymbol{Q}_i)$ guarantees that no $\boldsymbol{Q}_i$ will get rejected as an outlier. In fact, it implies that the inner minimization of (TLS-Q) always "chooses" the quadratic term $\boldsymbol{w}_0^\top \boldsymbol{Q}_i \boldsymbol{w}_0$ for any unit quaternion $\boldsymbol{w}_0$. While (SDR) "is aware of" $c_i^2 > \lambda_{\text{max}}(\boldsymbol{Q}_i)$ (e.g., when it holds, (SDR) is tight), this condition presents a gap from Theorem 1: $\lambda_{\text{max}}(\boldsymbol{Q}_i) \neq 0$ even in the absence of noise but Theorem 1 holds for every $c_i^2 > 0$. Thus, while Theorem 4 promises the tightness if c_i^2 is large enough, it leads us to the task of finding the smallest possible c_i^2 for which (SDR) remains tight. This turns out to be very challenging. In what follows, we give our efforts to this task, which we hope will provide further insights into the noisy and outlier-free case.

Smaller Truncation Parameters for The Tightness The smaller truncation parameters c_i^2 that we find are tightly related to the *eigengap* ζ, defined as the ratio between the second smallest eigenvalue $\lambda_{\text{min2}}(\cdot)$ of $\sum_{i=1}^{\ell} \boldsymbol{Q}_i$ and its minimum eigenvalue:

$$\zeta := \frac{\lambda_{\text{min2}}\big(\sum_{i=1}^{\ell} \boldsymbol{Q}_i\big)}{\lambda_{\text{min}}\big(\sum_{i=1}^{\ell} \boldsymbol{Q}_i\big)} \tag{24}$$

In analysis of eigenvalue algorithms (cf. [17]), the eigengap is typically defined as the *difference* between two consecutive eigenvalues of some matrix. Our eigengap (24) that takes the *division* of two smallest eigenvalues is not standard, but it will be convenient for our purpose. Also note that, $\boldsymbol{Q}_i$ is a perturbed version of $\boldsymbol{P}_i$ by noise $\boldsymbol{\epsilon}_i$ (Lemma 3), so $\lambda_{\text{min}}\big(\sum_{i=1}^{\ell}\boldsymbol{Q}_i\big)$ is in general nonzero, while it indeed approaches zero if $\boldsymbol{\epsilon}_i \to 0$. Clearly $\zeta \geq 1$. Moreover, we have the following immediate observation:

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

We are now ready to state the following result:

Theorem 5 (Noisy and Outlier-Free Point Sets, Version 2). *Suppose $\zeta \geq \ell/(\ell-1)$. The same conclusion of Theorem 4 holds true if*

$$c_i^2 > \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 + \big\|\boldsymbol{Q}_i \hat{\boldsymbol{w}}_0\big\|_2 + \frac{|d_i| + d_i}{2}, \quad \forall i = 1, \dots, \ell \tag{25}$$

$$\text{with } \; d_i := \frac{\sum_{i=1}^{\ell} \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0}{\ell} - \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 + \frac{\lambda_{\text{max}}\Big(\sum_{j\neq i}(\boldsymbol{Q}_i - \boldsymbol{Q}_j)\Big)}{\zeta(\ell-1)}. \tag{26}$$

Proof (Sketch). Let $\widehat{\boldsymbol{V}} := [\widehat{\boldsymbol{V}}_0, \;\; \hat{\boldsymbol{w}}_0] \in \mathbb{R}^{4\times 4}$ form an orthonormal basis of $\mathbb{R}^4$; $\widehat{\boldsymbol{V}}_0 \in \mathbb{R}^{4\times 3}$ satisfies $\widehat{\boldsymbol{V}}_0^\top \hat{\boldsymbol{w}}_0 = 0$ and $\widehat{\boldsymbol{V}}_0^\top \widehat{\boldsymbol{V}}_0 = \boldsymbol{I}_3$. Let $\hat{\mu} := \sum_{i=1}^{\ell}\big(\hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 - c_i^2\big)$ and

$$[\hat{\mathcal{D}}]_{0i} := \widehat{\boldsymbol{V}} \begin{bmatrix} \widehat{\boldsymbol{T}}_i & 0 \\ 0 & 0 \end{bmatrix} \widehat{\boldsymbol{V}}^\top - \frac{1}{2}\Big(\boldsymbol{Q}_i - c_i^2 \boldsymbol{I}_4\Big), \quad \forall i = 1, \dots, \ell \tag{27}$$

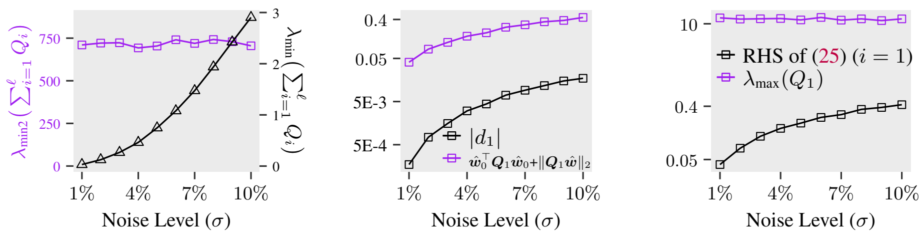
where $\widehat{\boldsymbol{T}}_i$ is a 3×3 symmetric matrix defined as

$$\widehat{\boldsymbol{T}}_i := \frac{\zeta - \frac{\ell}{\ell-1}}{\zeta} \widehat{\boldsymbol{V}}_0^\top \boldsymbol{Q}_i \widehat{\boldsymbol{V}}_0 + \frac{\sum_{j=1}^{\ell} \widehat{\boldsymbol{V}}_0^\top \boldsymbol{Q}_j \widehat{\boldsymbol{V}}_0}{\zeta(\ell-1)} - \frac{\Big(\sum_{i=1}^{\ell} \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0\Big)\boldsymbol{I}_3}{\ell}. \tag{28}$$

Then, similarly to the proof sketch of Theorem 4, it is not hard to show that conditions (O1) and (O3) of Proposition 2 are satisfied. Yet, proving (O2) under assumptions (25) and $\zeta \geq \ell/(\ell-1)$ is not that obvious, and we omit it here in interest of space.

Theorem 5 is better understood via numerics. We take randomly generated $\ell = 100$ point pairs $(\boldsymbol{y}_i, \boldsymbol{x}_i)$'s with $\boldsymbol{x}_i \sim \mathcal{N}(0, \boldsymbol{I}_3)$, and add different levels of Gaussian noise $\epsilon_i \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I}_3)$, where σ ranges from 1% to 10%. The values of $\lambda_{\text{min}}\big(\sum_{i=1}^{\ell}\boldsymbol{Q}_i\big)$ and $\lambda_{\text{min2}}\big(\sum_{i=1}^{\ell}\boldsymbol{Q}_i\big)$, and thus ζ, are shown in Figure 1a, where one might observe that $\zeta \approx 250$ for 10% noise, $\zeta \approx 25000$ 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 (25). First we recall that $c_i^2 > \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0$ is essential for (TLS-Q) to preserve all inliers. Second, we argue that the term $\|\boldsymbol{Q}_i \hat{\boldsymbol{w}}\|_2$ in (25) is also essential, as it accounts for the fact that noise destroys the inequality $\lambda_{\text{min}}(\boldsymbol{Q}_i) - \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 \geq 0$ which holds in the noiseless case (where $\lambda_{\text{min}}(\boldsymbol{Q}_i) = \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 = 0$) but gets violated (in general) in the presence of noise.



(a) (b) (c)

Fig. 1: Numerical illustration of condition (25) of Theorem 5 (500 trials, $\ell = 100$).

Finally, (25) also incurs a curious term $(|d_i| + d_i)/2$, with d_i defined in a sophisticated way (26). 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 $|d_1|$ in Figure 1b in comparison to $\hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_1 \hat{\boldsymbol{w}}_0 + \|\boldsymbol{Q}_1 \hat{\boldsymbol{w}}\|_2$, and observed that $|d_1|$ is two orders of magnitude smaller (there is nothing special about the choice of index 1). In fact, as noise approaches zero, we have $\boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 \to 0$ and (in general) $\zeta \to \infty$, hence $d_i \to 0$ by definition (26). Overall, condition (25) degenerates into $c_i^2 > 0$ in the noiseless case. Thus one might conclude that condition (25) is tighter than $c_i^2 > \lambda_{\text{max}}(\boldsymbol{Q}_i)$ of Theorem 4, as Lemma 1 implies $\lambda_{\text{max}}(\boldsymbol{Q}_i) \neq 0$ even without noise. Indeed, this is further numerically evidenced by Figure 1c where the lower bound of (25) with $i = 1$ ranges from 0.04 to 0.4, while the counterpart $\lambda_{\text{max}}(\boldsymbol{Q}_1)$ of Theorem 4 is roughly 12, arguably much larger.

While the term $(|d_i| + d_i)/2$ 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 $c_i^2 > \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 + \left\|\boldsymbol{Q}_i \hat{\boldsymbol{w}}_0\right\|_2$. 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.

3.4 The Noisy + Outliers Case

Combine the proof ideas of Theorems 2 and 5, and we obtain:

Theorem 6 (Noisy Point Sets with Outliers). *Let $\boldsymbol{Q}_1, \dots, \boldsymbol{Q}_{k^*}$ be inliers, the rest $\boldsymbol{Q}_j$'s outliers, and $\hat{\boldsymbol{w}}_0$ a global minimizer of* (TLS-Q). *Define*

$$\zeta_{\text{in}} := \frac{\lambda_{\text{min2}}\left(\sum_{i=1}^{k^*} \boldsymbol{Q}_i\right)}{\lambda_{\text{min}}\left(\sum_{i=1}^{k^*} \boldsymbol{Q}_i\right)}. \tag{29}$$

Assume (1) $\zeta_{\text{in}} \geq k^*/(k^* - 1)$, (2) for every $j = k^* + 1, \dots, \ell$, we have $0 < c_j^2 < \lambda_{\text{min}}(\boldsymbol{Q}_j)$, (3) for every $i = 1, \dots, k^*$, (25) *holds with d_i now defined as*

$$d_i := \frac{\sum_{i=1}^{k^*} \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0}{k^*} - \hat{\boldsymbol{w}}_0^\top \boldsymbol{Q}_i \hat{\boldsymbol{w}}_0 + \frac{\lambda_{\text{max}}\Big(\sum_{j \neq i} \left(\boldsymbol{Q}_i - \boldsymbol{Q}_j\right)\Big)}{\zeta(k^* - 1)}. \tag{30}$$

Then (SDR) *is tight and, similarly to* (23) *we have*

$$\sin^2(\hat{\tau}_0^*) \leq \frac{4\sum_{i=1}^{k^*}\|\boldsymbol{\epsilon}_i\|_2\|\boldsymbol{x}_i\|_2}{\lambda_{\text{min2}}\left(\sum_{i=1}^{k^*}\boldsymbol{P}_i\right)}, \quad \sin^2(\hat{\tau}_0^*) := 1-(\hat{\boldsymbol{w}}_0^\top \boldsymbol{w}_0^*)^2 \tag{31}$$

Here we recall that $\boldsymbol{w}_0^* \in \mathbb{S}^3$ *is the ground-truth unit quaternion, and each* $\boldsymbol{P}_i$ *is the "pure data" part of* $\boldsymbol{Q}_i$ *that satisfies* $\boldsymbol{P}_i\boldsymbol{w}_0^* = 0, \boldsymbol{P}_i \succeq 0$ *(Lemma 3).*

Proof. The given assumptions ensure that (TLS-Q) rejects all outliers and admit all inliers, and the minimum of (TLS-Q) is $\sum_{i=1}^{k^*}\hat{\boldsymbol{w}}_0^\top\boldsymbol{Q}_i\hat{\boldsymbol{w}}_0+\sum_{j=k^*+1}^{\ell}c_j^2$. For $i=1,\dots,k^*$ let $\mathcal{D}_{0i}$ be defined as in (27), and for $j=k^*+1,\dots,\ell$ let $[\hat{\mathcal{D}}]_{0j} := \frac{1}{2}(\boldsymbol{Q}_j - c_j^2\boldsymbol{I}_4)$. Let $\hat{\mu} := \sum_{i=1}^{k^*}\left(\hat{\boldsymbol{w}}_0^\top\boldsymbol{Q}_i\hat{\boldsymbol{w}}_0 - c_i^2\right)$. One then verifies the optimality conditions (O1) and (O3) of Proposition 2 are satisfied. (O2) is equivalent to ($\forall \boldsymbol{z}_i \in \mathbb{R}^4$)

$$\underbrace{-\hat{\mu}\|\boldsymbol{z}_0\|_2^2+\sum_{i=1}^{k^*}\left(2\boldsymbol{z}_i^\top[\hat{\mathcal{D}}]_{0i}\boldsymbol{z}_i-\boldsymbol{z}_0^\top\left(2[\hat{\mathcal{D}}]_{0i}-\boldsymbol{Q}_i+c_i^2\boldsymbol{I}_4\right)\boldsymbol{z}_i\right)}_{\text{Inlier Term}}+\underbrace{\sum_{i=k^*+1}^{\ell}O_j}_{\text{Outlier Term}} \geq 0,$$

where $O_j := \boldsymbol{z}_j(\boldsymbol{Q}_j - c_j^2\boldsymbol{I}_4)\boldsymbol{z}_j$. Since $\boldsymbol{Q}_j \succeq c_j^2\boldsymbol{I}_4$, O_j is non-negative. Under the given assumptions, one can replace ℓ by k^* in the proof of Theorem 5 and then find the inlier term is also non-negative. This finishes proving (O2) and thus the tightness of (SDR). The error bound (31) follows from the proof of Theorem 4 with ℓ replaced by k^*.

Note that all assumptions of Theorem 6 have their counterparts in previous results (e.g., Theorems 2 and 5), so we omitted further explanations.

4 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 [58] for 6 examples of truncated least-squares and see also [20,3,10].

As is common in the optimization literature, the relaxation we analyzed is at the first (i.e., lowest) relaxation order of the *Lasserre* hierarchy [36], or otherwise known as the *Shor relaxation* [50]. A tighter relaxation that has quadratically more constraints than (SDR) exists (cf. [57]). 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. [19,20]). Therefore, we leave this challenging question to future work.

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

References

1. Agostinho, S., Gomes, J., Del Bue, A.: CvxPnPL: A unified convex solution to the absolute pose estimation problem from point and line correspondences. Tech. rep., arXiv:1907.10545v2 [cs.CV] (2019)
2. Aholt, C., Agarwal, S., Thomas, R.: A QCQP approach to triangulation. In: European Conference on Computer Vision. pp. 654–667. Springer (2012)
3. Alfassi, Y., Keren, D., Reznick, B.: The non-tightness of a convex relaxation to rotation recovery. Sensors **21**(21), 7358 (2021)
4. Antonante, P., Tzoumas, V., Yang, H., Carlone, L.: Outlier-robust estimation: Hardness, minimally tuned algorithms, and applications. IEEE Transactions on Robotics (2021)
5. Bandeira, A.S., Boumal, N., Singer, A.: Tightness of the maximum likelihood semidefinite relaxation for angular synchronization. Mathematical Programming **163**(1-2), 145–167 (2017)
6. Boumal, N.: An introduction to optimization on smooth manifolds. Available online (2020), http://www.nicolasboumal.net/book
7. Briales, J., Gonzalez-Jimenez, J.: Fast global optimality verification in 3D SLAM. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 4630–4636 (2016)
8. Briales, J., Gonzalez-Jimenez, J.: Convex global 3D registration with lagrangian duality. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 4960–4969 (2017)
9. Briales, J., Kneip, L., Gonzalez-Jimenez, J.: A certifiably globally optimal solution to the non-minimal relative pose problem. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 145–154 (2018)
10. Brynte, L., Larsson, V., Iglesias, J.P., Olsson, C., Kahl, F.: On the tightness of semidefinite relaxations for rotation estimation. Journal of Mathematical Imaging and Vision **64**(1), 57–67 (2022)
11. Bustos, A.P., Chin, T.J.: Guaranteed outlier removal for rotation search. In: IEEE International Conference on Computer Vision. pp. 2165–2173 (2015)
12. Candes, E.J., Strohmer, T., Voroninski, V.: Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming. Communications on Pure and Applied Mathematics **66**(8), 1241–1274 (2013)
13. Carlone, L., Calafiore, G.C.: Convex relaxations for pose graph optimization with outliers. IEEE Robotics and Automation Letters **3**(2), 1160–1167 (2018)
14. Carlone, L., Calafiore, G.C., Tommolillo, C., Dellaert, F.: Planar pose graph optimization: Duality, optimal solutions, and verification. IEEE Transactions on Robotics **32**(3), 545–565 (2016)
15. Carlone, L., Rosen, D.M., Calafiore, G., Leonard, J.J., Dellaert, F.: Lagrangian duality in 3D SLAM: Verification techniques and optimal solutions. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 125–132 (2015)
16. Chaudhury, K.N., Khoo, Y., Singer, A.: Global registration of multiple point clouds using semidefinite programming. SIAM Journal on Optimization **25**(1), 468–501 (2015)
17. Chen, Y., Chi, Y., Fan, J., Ma, C., et al.: Spectral methods for data science: A statistical perspective. Foundations and Trends® in Machine Learning **14**(5), 566–806 (2021)
18. Cheng, Y., Lopez, J.A., Camps, O., Sznaier, M.: A convex optimization approach to robust fundamental matrix estimation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2170–2178 (2015)
19. Cifuentes, D.: A convex relaxation to compute the nearest structured rank deficient matrix. SIAM Journal on Matrix Analysis and Applications **42**(2), 708–729 (2021)

20. Cifuentes, D., Agarwal, S., Parrilo, P.A., Thomas, R.R.: On the local stability of semidefinite relaxations. Mathematical Programming pp. 1–35 (2021)
21. Ding, T., Yang, Y., Zhu, Z., Robinson, D.P., Vidal, R., Kneip, L., Tsakiris, M.C.: Robust homography estimation via dual principal component pursuit. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 6080–6089 (2020)
22. Elhamifar, E., Vidal, R.: Sparse subspace clustering: Algorithm, theory, and applications. IEEE transactions on pattern analysis and machine intelligence **35**(11), 2765–2781 (2013)
23. Eriksson, A., Olsson, C., Kahl, F., Chin, T.J.: Rotation averaging and strong duality. In: Conference on Computer Vision and Pattern Recognition. pp. 127–135 (2018)
24. Fredriksson, J., Olsson, C.: Simultaneous multiple rotation averaging using lagrangian duality. In: Asian Conference on Computer Vision. pp. 245–258. Springer (2012)
25. Garcia-Salguero, M., Briales, J., Gonzalez-Jimenez, J.: Certifiable relative pose estimation. Image and Vision Computing **109**, 104142 (2021)
26. Garcia-Salguero, M., Gonzalez-Jimenez, J.: Fast and robust certifiable estimation of the relative pose between two calibrated cameras. Journal of Mathematical Imaging and Vision **63**(8), 1036–1056 (2021)
27. Giamou, M., Ma, Z., Peretroukhin, V., Kelly, J.: Certifiably globally optimal extrinsic calibration from per-sensor egomotion. IEEE Robotics and Automation Letters **4**(2), 367–374 (2019)
28. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2004)
29. Horn, B.K.: Closed-form solution of absolute orientation using unit quaternions. Journal of the Optical Society of America A **4**(4), 629–642 (1987)
30. Iglesias, J.P., Olsson, C., Kahl, F.: Global optimality for point set registration using semidefinite programming. In: IEEE Conference on Computer Vision and Pattern Recognition (2020)
31. Ikami, D., Yamasaki, T., Aizawa, K.: Fast and robust estimation for unit-norm constrained linear fitting problems. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 8147–8155 (2018)
32. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: A review. ACM Computing Surveys **31**(3), 264–323 (1999)
33. Kahl, F., Henrion, D.: Globally optimal estimates for geometric reconstruction problems. International Journal of Computer Vision **74**(1), 3–15 (2007)
34. Khoo, Y., Kapoor, A.: Non-iterative rigid 2D/3D point-set registration using semidefinite programming. IEEE Transactions on Image Processing **25**(7), 2956–2970 (2016)
35. Lajoie, P.Y., Hu, S., Beltrame, G., Carlone, L.: Modeling perceptual aliasing in SLAM via discrete–continuous graphical models. IEEE Robotics and Automation Letters **4**(2), 1232–1239 (2019)
36. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM Journal on optimization **11**(3), 796–817 (2001)
37. Li, M., Liang, G., Luo, H., Qian, H., Lam, T.L.: Robot-to-robot relative pose estimation based on semidefinite relaxation optimization. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 4491–4498 (2020)
38. Ling, S.: Near-optimal bounds for generalized orthogonal procrustes problem via generalized power method. arXiv:2112.13725 [cs.IT] (2021)
39. Lu, C., Liu, Y.F., Zhang, W.Q., Zhang, S.: Tightness of a new and enhanced semidefinite relaxation for MIMO detection. SIAM Journal on Optimization **29**(1), 719–742 (2019)
40. Ma, Y., Soatto, S., Košecká, J., Sastry, S.: An Invitation to 3D Vision: From Images to Geometric Models, vol. 26. Springer (2004)
41. Maron, H., Dym, N., Kezurer, I., Kovalsky, S., Lipman, Y.: Point registration via efficient convex relaxation. ACM Transactions on Graphics **35**(4) (2016)

42. Ozyesil, O., Singer, A., Basri, R.: Stable camera motion estimation using convex programming. SIAM Journal on Imaging Sciences **8**(2), 1220–1262 (2015)
43. Parra Bustos, A., Chin, T.J.: Guaranteed outlier removal for point cloud registration with correspondences. IEEE Transactions on Pattern Analysis and Machine Intelligence **40**(12), 2868–2882 (2018)
44. Parra Bustos, A., Chin, T.J., Eriksson, A., Li, H., Suter, D.: Fast rotation search with stereographic projections for 3D registration. IEEE Transactions on Pattern Analysis and Machine Intelligence **38**(11), 2227–2240 (2016)
45. Peng, L., Fazlyab, M., Vidal, R.: Towards understanding the semidefinite relaxations of truncated least-squares in robust rotation search. Tech. rep., arXiv:2207.08350 [math.OC] (2022)
46. Peng, L., Tsakiris, M.C., Vidal, R.: ARCS: Accurate rotation and correspondences search. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11153–11163 (2022)
47. Probst, T., Paudel, D.P., Chhatkuli, A., Gool, L.V.: Convex relaxations for consensus and non-minimal problems in 3D vision. In: IEEE/CVF International Conference on Computer Vision. pp. 10233–10242 (2019)
48. Rosen, D.M., Carlone, L., Bandeira, A.S., Leonard, J.J.: SE-Sync: A certifiably correct algorithm for synchronization over the special euclidean group. The International Journal of Robotics Research **38**(2-3), 95–125 (2019)
49. Shi, J., Heng, Y., Carlone, L.: Optimal pose and shape estimation for category-level 3d object perception. In: Robotics: Science and Systems (2021)
50. Shor, N.Z.: Dual quadratic estimates in polynomial and boolean programming. Annals of Operations Research **25**(1), 163–168 (1990)
51. Tat-Jun, C., Zhipeng, C., Neumann, F.: Robust fitting in computer vision: Easy or hard? International Journal of Computer Vision **128**(3), 575–587 (2020)
52. Tian, Y., Khosoussi, K., Rosen, D.M., How, J.P.: Distributed certifiably correct pose-graph optimization. IEEE Transactions on Robotics **37**(6), 2137–2156 (2021)
53. Unnikrishnan, J., Haghighatshoar, S., Vetterli, M.: Unlabeled sensing with random linear measurements. IEEE Transactions on Information Theory **64**(5), 3237–3253 (2018)
54. Wang, L., Singer, A.: Exact and stable recovery of rotations for robust synchronization. Information and Inference: A Journal of the IMA **2**(2), 145–193 (2013)
55. Wise, E., Giamou, M., Khoubyarian, S., Grover, A., Kelly, J.: Certifiably optimal monocular hand-eye calibration. In: IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems. pp. 271–278. IEEE (2020)
56. Yang, H., Antonante, P., Tzoumas, V., Carlone, L.: Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection. IEEE Robotics and Automation Letters **5**(2), 1127–1134 (2020)
57. Yang, H., Carlone, L.: A quaternion-based certifiably optimal solution to the Wahba problem with outliers. In: IEEE International Conference on Computer Vision. pp. 1665–1674 (2019)
58. Yang, H., Carlone, L.: Certifiable outlier-robust geometric perception: Exact semidefinite relaxations and scalable global optimization. Tech. rep., arXiv:2109.03349 [cs.CV] (2021)
59. Yang, H., Shi, J., Carlone, L.: TEASER: Fast and certifiable point cloud registration. IEEE Transactions on Robotics **37**(2), 314–333 (2021)
60. Zhao, J.: An efficient solution to non-minimal case essential matrix estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence (2020)
61. Zhao, J., Xu, W., Kneip, L.: A certifiably globally optimal solution to generalized essential matrix estimation. In: IEEE Conference on Computer Vision and Pattern Recognition (2020)
62. Zhong, Y., Boumal, N.: Near-optimal bounds for phase synchronization. SIAM Journal on Optimization **28**(2), 989–1016 (2018)