

TEASER: Fast and Certifiable Point Cloud Registration

Heng Yang , Jingnan Shi , and Luca Carlone

Abstract—We propose the first fast and certifiable algorithm for the registration of two sets of three-dimensional (3-D) points in the presence of large amounts of outlier correspondences. A certifiable algorithm is one that attempts to solve an intractable optimization problem (e.g., robust estimation with outliers) and provides readily checkable conditions to verify if the returned solution is optimal (e.g., if the algorithm produced the most accurate estimate in the face of outliers) or bound its suboptimality or accuracy. Toward this goal, we first reformulate the registration problem using a truncated least squares (TLS) cost that makes the estimation insensitive to a large fraction of spurious correspondences. Then, we provide a general graph-theoretic framework to decouple scale, rotation, and translation estimation, which allows solving in cascade for the three transformations. Despite the fact that each subproblem (scale, rotation, and translation estimation) is still nonconvex and combinatorial in nature, we show that 1) TLS scale and (component-wise) translation estimation can be solved in polynomial time via an adaptive voting scheme, 2) TLS rotation estimation can be relaxed to a semidefinite program (SDP) and the relaxation is tight, even in the presence of extreme outlier rates, and 3) the graph-theoretic framework allows drastic pruning of outliers by finding the maximum clique. We name the resulting algorithm TEASER (*Truncated least squares Estimation And SEMidefinite Relaxation*). While solving large SDP relaxations is typically slow, we develop a second fast and certifiable algorithm, named TEASER++, that uses graduated nonconvexity to solve the rotation subproblem and leverages Douglas-Rachford Splitting to efficiently certify global optimality. For both algorithms, we provide theoretical bounds on the estimation errors, which are the first of their kind for robust registration problems. Moreover, we test their performance on standard benchmarks, object detection datasets, and the 3DMatch scan matching dataset, and show that 1) both algorithms dominate the state-of-the-art (e.g., RANSAC, branch-&-bound, heuristics) and are robust to more than 99% outliers when the scale is known, 2) TEASER++ can run in milliseconds and it is currently the fastest robust registration algorithm, and 3) TEASER++ is so robust it can also solve problems without correspondences (e.g., hypothesizing all-to-all correspondences).

Manuscript received January 23, 2020; revised August 8, 2020; accepted October 13, 2020. Date of publication December 8, 2020; date of current version April 2, 2021. This work was supported by ARL DCIST CRA W911NF-17-2-0181, ONR RAIDER N00014-18-1-2828, Lincoln Laboratory “Resilient Perception in Degraded Environments,” and the Google Daydream Research Program. This article was recommended for publication by Associate Editor J. M. M. Montiel and Editor P. Robuffo Giordano upon evaluation of the reviewers’ comments. (Corresponding author: Heng Yang.)

The authors are with the Laboratory for Information & Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: hkyang@mit.edu; jnshi@mit.edu; lcarlone@mit.edu).

This article has supplementary downloadable material available at <https://ieeexplore.ieee.org>, provided by the authors.

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>. Video: <https://youtu.be/xib1RSUoeQ> Code: <https://github.com/MIT-SPARK/TEASER-plusplus>.

Digital Object Identifier 10.1109/TRO.2020.3033695

where it largely outperforms ICP and it is more accurate than Go-ICP while being orders of magnitude faster. We release a fast open-source C++ implementation of TEASER++.

Index Terms—Certifiable algorithms, object pose estimation, outliers-robust estimation, point cloud alignment, robust estimation, scan matching, three-dimensional (3-D) registration, 3-D robot vision.

I. INTRODUCTION

Point cloud registration (also known as *scan matching* or *point cloud alignment*) is a fundamental problem in robotics and computer vision and consists in finding the best transformation (rotation, translation, and potentially scale) that aligns two point clouds. It finds applications in motion estimation and 3-D reconstruction [3]–[6], object recognition and localization [7]–[10], panorama stitching [11], and medical imaging [12], [13], to name a few.

When the ground-truth correspondences between the point clouds are known and the points are affected by zero-mean Gaussian noise, the registration problem can be readily solved, since elegant closed-form solutions [14], [15] exist for the case of isotropic noise. In practice, however, the correspondences are either unknown, or contain many outliers, leading these solvers to produce poor estimates. Large outlier rates are typical of 3-D keypoint detection and matching [16].

Commonly used approaches for registration with unknown or uncertain correspondences either rely on the availability of an initial guess for the unknown transformation (e.g., the iterative closest point, ICP [17]), or implicitly assume the presence of a small set of outliers (e.g., RANSAC [18]).¹ These algorithms may fail without notice [see Fig. 1 (b) and (e)] and may return estimates that are arbitrarily far from the ground-truth transformation. In general, the literature is divided between *heuristics*, that are fast but brittle, and *global methods* that are guaranteed to be robust, but run in worst-case exponential time (e.g., branch-&-bound methods, such as Go-ICP [20]).

This article is motivated by the goal of designing an approach that 1) can solve registration globally (without relying on an initial guess), 2) can tolerate extreme amounts of outliers (e.g., when 99% of the correspondences are outliers), 3) runs in polynomial time and is fast in practice (*i.e.*, can operate in real-time on a robot), and 4) provides formal performance guarantees. In particular, we look for *a posteriori* guarantees, e.g., conditions that one can check after executing the algorithm to assess the quality of the estimate. This leads to the notion of *certifiable algorithms*, *i.e.*, an algorithm that attempts to solve

¹RANSAC’s runtime grows exponentially with the outlier ratio [19] and it typically performs poorly with large outlier rates (see Sections II and XI).

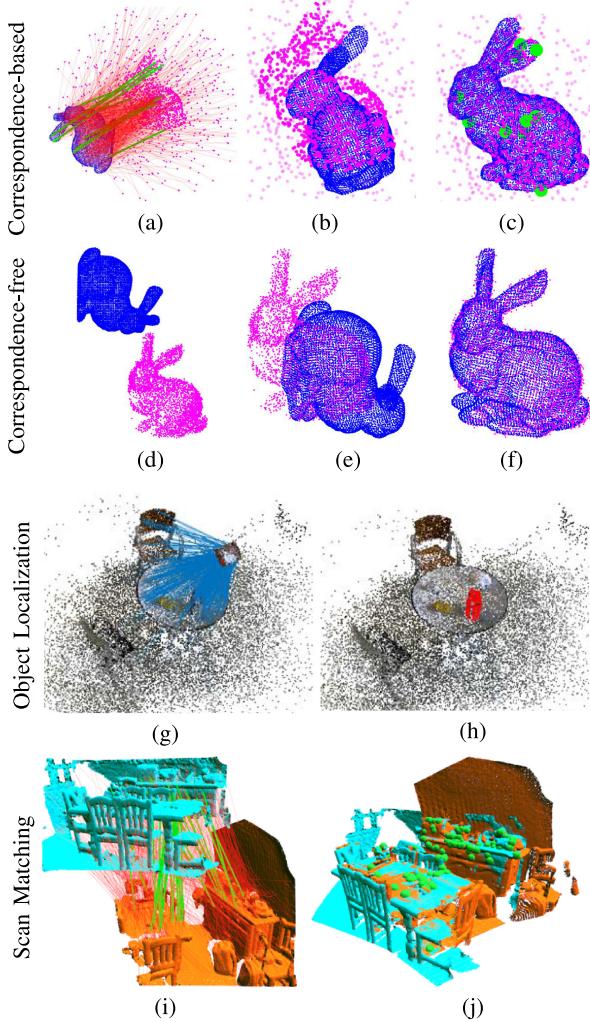


Fig. 1. We address 3-D registration in the realistic case where many point-to-point correspondences are outliers due to incorrect matching. (a) Bunny dataset (source cloud in blue and target cloud in magenta) with 95% outliers (shown as red lines) and 5% inliers (shown as green lines). Existing algorithms, such as RANSAC (b), can produce incorrect estimates without notice even after running for 10000 iterations. Our certifiable algorithm, TEASER, largely outperforms the state of the art in terms of robustness and accuracy, and a fast implementation, named TEASER++ (c), computes accurate estimates in milliseconds even with extreme outlier rates and finds the small set of inliers (shown as green dots). The unprecedented robustness of TEASER++ enables the solution of correspondence-free registration problems (d), where ICP (e) fails without a good initial guess, while TEASER++ (f) succeeds without requiring an initial guess. We test our approach in challenging object localization (g-h) and scan matching (i-j) RGB-D datasets, using both traditional features (*i.e.*, FPFH [1]) and deep-learned features (*i.e.*, 3DSmoothNet [2]). (a) Input. (b) RANSAC. (c) TEASER++. (d) Input. (e) ICP. (f) TEASER++. (g) Correspondences. (h) TEASER++. (i) Correspondences. (j) TEASER++.

如果存在大量的异常(匹配),在实际进行配准时应该使用的是什么样的算法,为什么能够处理异常值?

an intractable problem and provides checkable conditions on whether it succeeded [21]–[26]. The interested reader can find a broader discussion in Appendix A in [27].

Contribution: This article proposes the first certifiable algorithm for 3-D registration with outliers. We reformulate the registration problem using a truncated least squares (TLS) cost (presented in Section IV), which is insensitive to a large fraction of spurious correspondences but leads to a hard, combinatorial, and nonconvex optimization.

The first contribution (see Section V) is a general framework to decouple scale, rotation, and translation estimation. The idea of decoupling rotation and translation has appeared in related work, e.g., [19], [28], [29]. The novelty of our proposal is fourfold: (i) we develop invariant measurements to estimate the scale ([19], [29] assume the scale is given), (ii) we make the decoupling formal under the assumption of *unknown-but-bounded noise* [30], [31], (iii) we provide a general graph-theoretic framework to derive these invariant measurements, and (iv) we show that this framework allows pruning a large amount of outliers by finding the *maximum clique* of the graph defined by the invariant measurements (see Section VI).

The decoupling allows solving in cascade for scale, rotation, and translation. However, each subproblem is still combinatorial in nature. Our second contribution is to show that 1) in the scalar case TLS estimation can be solved exactly in polynomial time using an *adaptive voting* scheme, and this enables efficient estimation of the scale and the (component-wise) translation (see Section VII); 2) we can formulate a tight *semidefinite programming (SDP) relaxation* to estimate the rotation and establish *a posteriori* conditions to check the quality of the relaxation (see Section VIII). We remark that the rotation subproblem addressed in this article is in itself a foundational problem in vision (where it is known as *rotation search* [32]) and aerospace (where it is known as the *Wahba problem* [33]). Our SDP relaxation is the first certifiable algorithm for robust rotation search.

Our third contribution (see Section IX) is a set of theoretical results certifying the quality of the solution returned by our algorithm, named *Truncated least squares Estimation And SEMidefinite Relaxation* (TEASER). In the noiseless case, we provide easy-to-check conditions under which TEASER recovers the true transformation between the point clouds in the presence of outliers. In the noisy case, we provide bounds on the distance between the ground-truth transformation and TEASER's estimate. To the best of our knowledge these are the first nonasymptotic error bounds for geometric estimation problems with outliers, while the literature on robust estimation in statistics (e.g., [34]) typically studies simpler problems in Euclidean space and focuses on asymptotic bounds.

Our fourth contribution (see Section X) is to implement a fast version of TEASER, named TEASER++, that uses *graduated nonconvexity* (GNC) [35] to estimate the rotation without solving a large SDP. We show that TEASER++ is also certifiable, and in particular we leverage *Douglas–Rachford Splitting (DRS)* [36] to design a scalable *optimality certifier* that can assert global optimality of the estimate returned by GNC. We release a fast open-source C++ implementation of TEASER++.

Our last contribution (see Section XI) is an extensive evaluation in both standard benchmarks and on real datasets for object detection [37] and scan matching [38]. In particular, we show that 1) both TEASER and TEASER++ dominate the state of the art (e.g., RANSAC, branch-&-bound, heuristics) and are robust to more than 99% outliers when the scale is known, 2) TEASER++ can run in milliseconds and it is currently the fastest robust registration algorithm, 3) TEASER++ is so robust it can also solve problems without correspondences (e.g., hypothesizing all-to-all correspondences), where it largely outperforms ICP and it is more accurate than Go-ICP while being orders of magnitude faster, and 4) TEASER++ can boost registration performance when combined with deep-learned keypoint detection and matching.

All the appendices with proofs, extra experimental results, and discussions are given in the Supplementary Material [27].

Novelty with respect to [39], [40]: In our previous works, we introduced TEASER [39] and the quaternion-based relaxation of the rotation subproblem [40] (named QUASAR). The present manuscript brings TEASER to maturity by 1) providing explicit theoretical results on TEASER's performance (see Section IX), 2) providing a fast optimality certification method (see Section VIII-C), 3) developing a fast algorithm, TEASER++, that uses GNC to estimate the rotation without solving an SDP, while still being certifiable (see Section X), and 4) reporting a more comprehensive experimental evaluation, including real tests on the 3DMatch dataset and examples of registration without correspondences (see Sections XI-C and XI-E). These are major improvements both on the theoretical side ([39], [40] only certified performance on each subproblem, rather than end-to-end, and required solving a large SDP) and on the practical side (TEASER++ is more than three orders of magnitude faster than our proposal in [39]).

II. RELATED WORK

There are two popular paradigms for the registration of 3-D point clouds: *Correspondence-based* and *Simultaneous Pose and Correspondence* (SPC) (*i.e.*, *correspondence-free*) methods.

A. Correspondence-Based Methods

Correspondence-based methods first detect and match 3-D keypoints between point clouds using feature descriptors [1], [7], [16], [41], and then use an estimator to infer the transformation from these putative correspondences. *3-D keypoint matching is known to be less accurate compared to 2-D counterparts like SIFT and ORB*, thus causing much higher outlier rates, *e.g.*, having 95% spurious correspondences is considered common [19]. Therefore, a robust backend that can deal with extreme outlier rates is highly desirable. 作为后端

Registration Without Outliers: Horn [14] and Arun [15] show that optimal solutions (in the maximum likelihood sense) for scale, rotation, and translation can be computed in closed form when the correspondences are known and the points are affected by isotropic zero-mean Gaussian noise. Olsson *et al.* [42] propose a method based on Branch-&-bound (BnB) that is globally optimal and allows point-to-point, point-to-line, and point-to-plane correspondences. Briales and Gonzalez-Jimenez [43] proposed a tight semidefinite relaxation to solve the same registration problem as in [42].

If the two sets of points only differ by an unknown rotation (*i.e.*, we do not attempt to estimate the scale and translation), we obtain a simplified version of the registration problem that is known as *rotation search* in computer vision [32], or *Wahba problem* in aerospace [33]. In aerospace, the vector observations are typically the directions to visible stars observed by sensors onboard the satellite. Closed-form solutions to the Wahba problem are known using both quaternion [14], [44] and rotation matrix [15], [45] representations. The Wahba problem is also related to the well-known *Orthogonal Procrustes problem* [46] where one searches for orthogonal matrices (rather than rotations), for which a closed-form solution also exists [47]. The computer vision community has investigated the rotation search problem in the context of point cloud registration [17], [20], image

stitching [11], motion estimation and 3-D reconstruction [4], [5]. In particular, the closed-form solutions from Horn [14] and Arun *et al.* [15] can be used for (outlier-free) rotation search with isotropic Gaussian noise. Ohta and Kanatani [48] proposed a quaternion-based optimal solution, and Cheng and Crassidis [49] developed a local optimization algorithm, for the case of anisotropic Gaussian noise. Ahmed *et al.* [50] developed an SDP relaxation for the case with bounded noise and no outliers.

Robust Registration: Probably the most widely used robust registration approach is based on RANSAC [18], which has enabled several early applications in vision and robotics [51], [52]. Despite its efficiency in the low-noise and low-outlier regime, RANSAC exhibits slow convergence and low accuracy with large outlier rates [19], where it becomes harder to sample a “good” consensus set. Other approaches resort to *M-estimation*, which replaces the least squares objective function with robust costs that are less sensitive to outliers [53]–[55]. Zhou *et al.* [56] proposed *fast global registration* (FGR) that uses the Geman-McClure cost function and leverages GNC to solve the resulting nonconvex optimization. Despite its efficiency, FGR offers no optimality guarantees. Indeed, FGR tends to fail when the outlier ratio is high (>80%), as we show in Section XI. Enqvist *et al.* [57] proposed to optimally find correct correspondences by solving the vertex cover problem and demonstrate robustness against over 95% outliers in 3-D registration. Parra and Chin [19] proposed a *Guaranteed Outlier Removal* (GORE) technique, that uses geometric operations to significantly reduce the amount of outlier correspondences before passing them to the optimization backend. GORE has been shown to be robust to 95% spurious correspondences [19]. In an independent effort, Parra *et al.* [58] found pairwise-consistent correspondences in 3-D registration using a practical maximum clique (PMC) algorithm. While similar in spirit to our original proposal [39], both GORE and PMC do not estimate the *scale* of the registration and are typically slower since they rely on BnB (see [19, Algorithm 2]). Yang and Carbone [39] proposed the first certifiable algorithm for robust registration, which however requires solving a large-scale SDP hence being limited to small problems. Recently, physics-based registration has been proposed in [59], [60], and [61].

Similar to the registration problem, rotation search with outliers has also been investigated in the computer vision community. Local techniques for robust rotation search are again based on RANSAC or M-estimation, but are brittle and do not provide performance guarantees. Global methods (that guarantee to compute globally optimal solutions) are based on *Consensus Maximization* [62]–[67] and BnB [68]. Hartley and Kahl [32] first proposed using BnB for rotation search, and Bazin *et al.* [69] adopted consensus maximization to extend their BnB algorithm with a robust formulation. BnB is guaranteed to return the globally optimal solution, but it runs in exponential time in the worst case. Another class of global methods for consensus maximization enumerates all possible subsets of measurements with size no larger than the problem dimension (three for rotation search) to analytically compute candidate solutions, and then verify global optimality using computational geometry [70], [71]. Similarly, Ask *et al.* [72] showed the TLS estimation can also be solved globally by enumerating subsets with size no larger than the model dimension. However, these methods require exhaustive enumeration and become intractable when the problem dimension is large (*e.g.*, 6 for 3-D registration). In [40],

问题的
核心其实
在于将其
进行分
解？

单纯的

rotation的问
题是有一个
封闭解的，
解也就是对
解析解，对
vision community
应于数值解
has investigated
the rotation search
problem in the
context of point
cloud registration
[17], [20], image

用BnB的方
法就一定慢
么？

关于全局的
方法还可以
再研究一下

we proposed the first quaternion-based certifiable algorithm for robust rotation search (reviewed in Section VIII).

B. Simultaneous Pose and Correspondence Methods

SPC methods alternate between finding the correspondences and computing the best transformation given the correspondences.

Local Methods: The *Iterative closest point* (ICP) algorithm [17] is considered a milestone in point cloud registration. However, ICP is prone to converge to *local minima* and only performs well given a good initial guess. Multiple variants of ICP [73]–[76] have proposed to use robust cost functions to improve convergence. Probabilistic interpretations have also been proposed to improve ICP as a minimization of the Kullback–Leibler divergence between two mixture models [77], [78]. Clark *et al.* [79] align point clouds as continuous functions using Riemannian optimization. Le *et al.* [80] used semidefinite relaxation to generate hypotheses for randomized methods. All these methods do not provide global optimality guarantees.

Global Methods: Global SPC approaches compute a globally optimal solution without initial guesses, and are usually based on BnB. A series of geometric techniques have been proposed to improve the bounding tightness [20], [32], [81], [82] and increase the search speed [20], [83]. However, the runtime of BnB increases exponentially with the size of the point cloud and is made worse by the explosion of the number of local minima resulting from high outlier ratios [19]. Global SPC registration can be also formulated as a mixed-integer program [84], though the runtime remains exponential. Maron *et al.* [85] designed a tight semidefinite relaxation, which, however, only works for SPC registration with full overlap.

Deep Learning Methods: The success of deep learning on 3-D point clouds (e.g., PointNet [86] and DGCNN [87]) opens new opportunities for learning point cloud registration from data. Deep learning methods first learn to embed points clouds in a high-dimensional feature space, then learn to match keypoints to generate correspondences, after which optimization over the space of rigid transformations is performed for the best alignment. PointNetLK [88] uses PointNet to learn feature representations and then iteratively align the features representations, instead of the 3-D coordinates. DCP [89] uses DGCNN features for correspondence matching and Horn’s method for registration in an end-to-end fashion (Horn’s method is differentiable). PRNet [90] extends DCP to aligning partially overlapping point clouds. Scan2CAD [91] and its improvement [92] apply similar pipelines to align CAD models to RGB-D scans. 3DSmoothNet [2] uses a siamese deep learning architecture to establish keypoint correspondences between two point clouds. FCGF [41] leverages sparse high-dimensional convolutions to extract dense feature descriptors from point clouds. Deep global registration [93] uses FCGF feature descriptors for point cloud registration. In Section XI-E, we show that 1) our approach provides a robust back-end for deep-learned keypoint matching algorithms (we use [2]), and 2) that current deep learning approaches still struggle to produce acceptable inlier rates in real problems.

Remark 1 (Reconciling Correspondence-based and SPC Methods): Correspondence-based and SPC methods are tightly coupled. First of all, approaches like ICP alternate between finding the correspondences and solving a correspondence-base problem. More importantly, one can always reformulate an SPC

problem as a correspondence-based problem by hypothesizing all-to-all correspondences, *i.e.*, associating each point in the first point cloud to all the points in the second. To the best of our knowledge, only [57], [94] have pursued this formulation since it leads to an extreme number of outliers. In this article, we show that our approach can indeed solve the SPC problem thanks to its unprecedented robustness to outliers (see Section XI-C).

III. NOTATIONS AND PRELIMINARIES

Scalars, Vectors, Matrices: We use lowercase characters (e.g., s) to denote real scalars, bold lowercase characters (e.g., \mathbf{v}) for real vectors, and bold uppercase characters (e.g., \mathbf{M}) for real matrices. M_{ij} denotes the i th row and j th column scalar entry of matrix $\mathbf{M} \in \mathbb{R}^{m \times n}$, and $[\mathbf{M}]_{ij,d}$ (or simply $[\mathbf{M}]_{ij}$ when d is clear from the context) denotes the i th row and j th column $d \times d$ block of the matrix $\mathbf{M} \in \mathbb{R}^{md \times nd}$. \mathbf{I}_d is the identity matrix of size d . We use “ \otimes ” to denote the Kronecker product. For a square matrix \mathbf{M} , $\det(\mathbf{M})$ and $\text{tr}(\mathbf{M})$ denote its determinant and trace. The two-norm of a vector is denoted as $\|\cdot\|$. The Frobenius norm of a matrix is denoted as $\|\cdot\|_F$. For a symmetric matrix \mathbf{M} of size $n \times n$, we use $\lambda_1 \leq \dots \leq \lambda_n$ to denote its real eigenvalues. 范数

Sets: We use calligraphic fonts to denote sets (e.g., \mathcal{S}). We use \mathcal{S}^n (resp. \mathfrak{S}^n) to denote the group of real symmetric (resp. skew-symmetric) matrices with size $n \times n$. $\mathcal{S}_+^n \doteq \{\mathbf{M} \in \mathcal{S}^n : \mathbf{M} \succeq 0\}$ denotes the set of $n \times n$ symmetric positive semidefinite (PSD) matrices. $\text{SO}(d) \doteq \{\mathbf{R} \in \mathbb{R}^{d \times d} : \mathbf{R}^\top \mathbf{R} = \mathbf{I}_d, \det(\mathbf{R}) = +1\}$ denotes the d -dimensional special orthogonal group, while $\mathbb{S}^{d-1} = \{\mathbf{u} \in \mathbb{R}^d : \|\mathbf{u}\| = 1\}$ denotes the d -dimensional unit sphere.

Quaternions: Unit quaternions are a representation for a 3-D rotation $\mathbf{R} \in \text{SO}(3)$. We denote a unit quaternion as a unit-norm column vector $\mathbf{q} = [\mathbf{v}^\top \ s]^\top \in \mathbb{S}^3$, where $\mathbf{v} \in \mathbb{R}^3$ is the *vector part* of the quaternion and the last element s is the *scalar part*. We also use $\mathbf{q} = [q_1 \ q_2 \ q_3 \ q_4]^\top$ to denote the four entries of the quaternion. Each quaternion represents a 3-D rotation and the composition of two rotations \mathbf{q}_a and \mathbf{q}_b can be computed using the *quaternion product* $\mathbf{q}_c = \mathbf{q}_a \circ \mathbf{q}_b$:

$$\mathbf{q}_c = \mathbf{q}_a \circ \mathbf{q}_b = \Omega_1(\mathbf{q}_a)\mathbf{q}_b = \Omega_2(\mathbf{q}_b)\mathbf{q}_a \quad (1)$$

where $\Omega_1(\mathbf{q})$ and $\Omega_2(\mathbf{q})$ are defined as follows:

$$\Omega_1(\mathbf{q}) = \begin{bmatrix} q_4 & -q_3 & q_2 & q_1 \\ q_3 & q_4 & -q_1 & q_2 \\ -q_2 & q_1 & q_4 & q_3 \\ -q_1 & -q_2 & -q_3 & q_4 \end{bmatrix}, \quad \Omega_2(\mathbf{q}) = \begin{bmatrix} q_4 & q_3 & -q_2 & q_1 \\ -q_3 & q_4 & q_1 & q_2 \\ q_2 & -q_1 & q_4 & q_3 \\ -q_1 & -q_2 & -q_3 & q_4 \end{bmatrix}. \quad (2)$$

The inverse of a quaternion $\mathbf{q} = [\mathbf{v}^\top \ s]^\top$ is defined as $\mathbf{q}^{-1} = [-\mathbf{v}^\top \ s]^\top$, where one simply reverses the sign of the vector part. The rotation of a vector $\mathbf{a} \in \mathbb{R}^3$ can be expressed in terms of quaternion product. Formally, if \mathbf{R} is the (unique) rotation matrix corresponding to a unit quaternion \mathbf{q} , then

$$\begin{bmatrix} \mathbf{Ra} \\ 0 \end{bmatrix} = \mathbf{q} \circ \hat{\mathbf{a}} \circ \mathbf{q}^{-1} \quad (3)$$

where $\hat{\mathbf{a}} = [\mathbf{a}^\top \ 0]^\top$ is the *homogenization* of \mathbf{a} , obtained by augmenting \mathbf{a} with an extra entry equal to zero. The set of unit quaternions, *i.e.*, the four-dimensional unit sphere \mathbb{S}^3 , is a *double cover* of $\text{SO}(3)$ since \mathbf{q} and $-\mathbf{q}$ represent the same rotation [this fact can be easily seen by examining (3)].

IV. ROBUST REGISTRATION WITH TRUNCATED LEAST SQUARES COST

In the robust registration problem, we are given two 3-D point clouds $\mathcal{A} = \{\mathbf{a}_i\}_{i=1}^N$ and $\mathcal{B} = \{\mathbf{b}_i\}_{i=1}^N$, with $\mathbf{a}_i, \mathbf{b}_i \in \mathbb{R}^3$. We consider a correspondence-based setup, where we are given putative correspondences $(\mathbf{a}_i, \mathbf{b}_i), i = 1, \dots, N$, that obey the following generative model:

$$\mathbf{b}_i = s^\circ \mathbf{R}^\circ \mathbf{a}_i + \mathbf{t}^\circ + \mathbf{o}_i + \boldsymbol{\epsilon}_i \quad (4)$$

where $s^\circ > 0$, $\mathbf{R}^\circ \in \text{SO}(3)$, and $\mathbf{t}^\circ \in \mathbb{R}^3$ are the unknown (to-be-computed) scale, rotation, and translation, $\boldsymbol{\epsilon}_i$ models the measurement noise, and \mathbf{o}_i is a vector of zeros if the pair $(\mathbf{a}_i, \mathbf{b}_i)$ is an *inlier*, or a vector of arbitrary numbers for *outlier correspondences*. In words, if the i th correspondence $(\mathbf{a}_i, \mathbf{b}_i)$ is an inlier correspondence, \mathbf{b}_i corresponds to a 3-D transformation of \mathbf{a}_i (plus noise $\boldsymbol{\epsilon}_i$), while if $(\mathbf{a}_i, \mathbf{b}_i)$ is an outlier correspondence, \mathbf{b}_i is just an arbitrary vector.

Registration Without Outliers: When $\boldsymbol{\epsilon}_i$ is a zero-mean Gaussian noise with isotropic covariance $\sigma_i^2 \mathbf{I}_3$, and all the correspondences are correct (i.e., $\mathbf{o}_i = \mathbf{0}, \forall i$), the maximum likelihood estimator of $(s^\circ, \mathbf{R}^\circ, \mathbf{t}^\circ)$ can be computed by solving the following nonlinear least squares problem:

$$\min_{s>0, \mathbf{R} \in \text{SO}(3), \mathbf{t} \in \mathbb{R}^3} \sum_{i=1}^N \frac{1}{\sigma_i^2} \|\mathbf{b}_i - s\mathbf{R}\mathbf{a}_i - \mathbf{t}\|^2. \quad (5)$$

Although (5) is a nonconvex problem, due to the nonconvexity of the set $\text{SO}(3)$, its optimal solution can be computed in closed form by decoupling the estimation of the scale, rotation, and translation, using Horn's [14] or Arun's method [15]. A key contribution of this present article is to provide a way to decouple scale, rotation, and translation in the more challenging case with outliers.

In practice, a large fraction of the correspondences are *outliers*, due to incorrect keypoint matching. Despite the elegance of the closed-form solutions [14], [15], they are not robust to outliers, and a single “bad” outlier can compromise the correctness of the resulting estimate. Hence, we propose a *truncated least squares registration* formulation that can tolerate extreme amounts of spurious data.

Truncated Least Squares Registration: We depart from the Gaussian noise model and assume the noise is *unknown but bounded* [30]. Formally, we assume the *inlier* noise $\boldsymbol{\epsilon}_i$ in (4) is such that $\|\boldsymbol{\epsilon}_i\| \leq \beta_i$, where β_i is a given bound.

Then we adopt the following TLS registration formulation:

$$\min_{s>0, \mathbf{R} \in \text{SO}(3), \mathbf{t} \in \mathbb{R}^3} \sum_{i=1}^N \min \left(\frac{1}{\beta_i^2} \|\mathbf{b}_i - s\mathbf{R}\mathbf{a}_i - \mathbf{t}\|^2, \bar{c}^2 \right) \quad (6)$$

which computes a least squares solution of measurements with small residuals ($\frac{1}{\beta_i^2} \|\mathbf{b}_i - s\mathbf{R}\mathbf{a}_i - \mathbf{t}\|^2 \leq \bar{c}^2$), while discarding measurements with large residuals (when $\frac{1}{\beta_i^2} \|\mathbf{b}_i - s\mathbf{R}\mathbf{a}_i - \mathbf{t}\|^2 > \bar{c}^2$ the i th summand becomes a constant and does not influence the optimization). Note that one can always divide each summand in (6) by \bar{c}^2 : therefore, one can safely assume \bar{c}^2 to be 1. For the sake of generality, in the following we keep \bar{c}^2 , since it provides a more direct “knob” to be stricter or more lenient toward potential outliers.

The noise bound β_i is fairly easy to set in practice and can be understood as a “3-sigma” noise bound or as the maximum error

we expect from an inlier. The interested reader can find a more formal discussion on how to set β_i and \bar{c} in Appendix B in [27]. We remark that while we assume to have a bound on the maximum error we expect from the inliers (β_i), we do not make assumptions on the generative model for the outliers, which is typically unknown in practice.

Remark 2 (TLS versus Consensus Maximization): TLS estimation is related to *Consensus Maximization* [62], a popular robust estimation approach in computer vision. Consensus Maximization looks for an estimate that maximizes the number of inliers, while TLS simultaneously computes a least squares estimate for the inliers. The two methods are not guaranteed to produce the same choice of inliers in general, since TLS also penalizes inliers with large errors. Appendix C in [27] provides a toy example to illustrate the potential mismatch between the two techniques and provides necessary conditions under which the two formulations find the same set of inliers.

Despite being insensitive to outlier correspondences, the truncated least squares formulation (6) is much more challenging to solve globally, compared to the outlier-free case (5). This is the case even in simpler estimation problems, where the feasible set of the unknowns is convex, as stated below.

Remark 3 (Hardness of TLS Estimation [95], [96]): The minimization of a sum of truncated convex functions, $\sum_{i=1}^N \min(f_i(x), \bar{c}^2)$ over a convex feasible set $x \in \mathcal{X} \subseteq \mathbb{R}^d$ is NP-hard in the dimension d [95], [96]. A simple exhaustive search can obtain a globally optimal solution in exponential time $O(2^N)$ [96].

Remark 3 states the NP-hardness of TLS estimation over a convex feasible set. The TLS problem (6) is even more challenging due to the nonconvexity of $\text{SO}(3)$. While problem (6) is hard to solve directly, in the next section, we show how to decouple the estimation of scale, rotation, and translation using *invariant measurements*.

V. DECOUPLING SCALE, ROTATION, AND TRANSLATION ESTIMATION

We propose a general approach to decouple the estimation of scale, translation, and rotation in problem (6). The key insight is that we can reformulate the measurements (4) to obtain quantities that are invariant to a subset of the transformations (scaling, rotation, and translation).

A. Translation Invariant Measurements (TIMs)

While the absolute positions of the points in \mathcal{B} depend on the translation \mathbf{t} , the relative positions are invariant to \mathbf{t} . Mathematically, given two points \mathbf{b}_i and \mathbf{b}_j from (4), the relative position of these two points is

$$\mathbf{b}_j - \mathbf{b}_i = s\mathbf{R}(\mathbf{a}_j - \mathbf{a}_i) + (\mathbf{o}_j - \mathbf{o}_i) + (\boldsymbol{\epsilon}_j - \boldsymbol{\epsilon}_i) \quad (7)$$

where the translation \mathbf{t} cancels out in the subtraction. Therefore, we can obtain a TIM by computing $\bar{\mathbf{a}}_{ij} \doteq \mathbf{a}_j - \mathbf{a}_i$ and $\bar{\mathbf{b}}_{ij} \doteq \mathbf{b}_j - \mathbf{b}_i$, and the TIM satisfies the following generative model:

$$\text{对于同一个圆形而言, 真两个点之间的相对距离是不变的} \quad \bar{\mathbf{b}}_{ij} = s\mathbf{R}\bar{\mathbf{a}}_{ij} + \mathbf{o}_{ij} + \boldsymbol{\epsilon}_{ij} \quad (\text{TIM})$$

where $\mathbf{o}_{ij} \doteq \mathbf{o}_j - \mathbf{o}_i$ is zero if both the i th and the j th measurements are inliers (or arbitrary otherwise), while $\boldsymbol{\epsilon}_{ij} \doteq \boldsymbol{\epsilon}_j - \boldsymbol{\epsilon}_i$ is the measurement noise. It is easy to see that if $\|\boldsymbol{\epsilon}_i\| \leq \beta_i$ and $\|\boldsymbol{\epsilon}_j\| \leq \beta_j$, then $\|\boldsymbol{\epsilon}_{ij}\| \leq \beta_i + \beta_j \doteq \delta_{ij}$.

各向同性的高斯分布就是球形的高斯分布

如何解决 outliers 的问题?

这个和 rejecting points 有什么区别么?

似乎没有发现什么好处

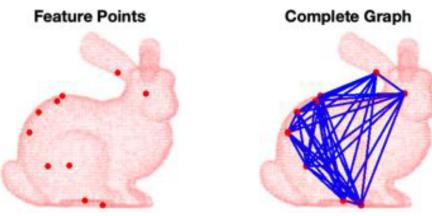


Fig. 2. TIMs generated from a complete graph in the Bunny dataset [98].
将两个点之间的距离作为参数

The advantage of the TIMs in eq. (TIM) is that their generative model only depends on two unknowns, s and \mathbf{R} . The number of TIMs is upper-bounded by $\binom{N}{2} = N(N - 1)/2$, where pairwise relative measurements between all pairs of points are computed. Theorem 4 below connects the TIMs with the topology of a graph defined over the 3-D points.

Theorem 4 (Translation Invariant Measurements): Define the vectors $\mathbf{a} \in \mathbb{R}^{3N}$ (resp. $\mathbf{b} \in \mathbb{R}^{3N}$), obtained by concatenating all vectors \mathbf{a}_i (resp. \mathbf{b}_i) in a single column vector. Moreover, define an arbitrary graph \mathcal{G} with nodes $\{1, \dots, N\}$ and an arbitrary set of edges \mathcal{E} . Then, the vectors $\bar{\mathbf{a}} = (\mathbf{A} \otimes \mathbf{I}_3)\mathbf{a}$ and $\bar{\mathbf{b}} = (\mathbf{A} \otimes \mathbf{I}_3)\mathbf{b}$ are TIMs, where $\mathbf{A} \in \mathbb{R}^{|\mathcal{E}| \times N}$ is the incidence matrix of \mathcal{G} [97].

A proof of the theorem is given in Appendix D in [27]. TIMs generated from a complete graph on the Bunny dataset [98] are illustrated in Fig. 2.

B. Translation and Rotation Invariant Measurements (TRIMs)

While the relative locations of pairs of points (TIMs) still depend on the rotation \mathbf{R} , their distances are invariant to both \mathbf{R} and \mathbf{t} . Therefore, to build rotation invariant measurements, we compute the norm of each TIM vector

$$\|\bar{\mathbf{b}}_{ij}\| = \|s\mathbf{R}\bar{\mathbf{a}}_{ij} + \mathbf{o}_{ij} + \epsilon_{ij}\|. \quad (8)$$

We now note that for the inliers ($\mathbf{o}_{ij} = \mathbf{0}$) it holds (using $\|\epsilon_{ij}\| \leq \delta_{ij}$ and the triangle inequality)

$$\|s\mathbf{R}\bar{\mathbf{a}}_{ij}\| - \delta_{ij} \leq \|s\mathbf{R}\bar{\mathbf{a}}_{ij} + \epsilon_{ij}\| \leq \|s\mathbf{R}\bar{\mathbf{a}}_{ij}\| + \delta_{ij} \quad (9)$$

hence we can write (8) equivalently as

$$\|\bar{\mathbf{b}}_{ij}\| = \|s\mathbf{R}\bar{\mathbf{a}}_{ij}\| + \tilde{o}_{ij} + \tilde{\epsilon}_{ij} \quad (10)$$

with $|\tilde{\epsilon}_{ij}| \leq \delta_{ij}$, and $\tilde{o}_{ij} = 0$ if both i and j are inliers or is an arbitrary scalar otherwise. Recalling that the norm is rotation invariant and that $s > 0$, and dividing both sides of (10) by $\|\bar{\mathbf{a}}_{ij}\|$, we obtain new measurements $s_{ij} \doteq \frac{\|\bar{\mathbf{b}}_{ij}\|}{\|\bar{\mathbf{a}}_{ij}\|}$:

$$s_{ij} = s + o_{ij}^s + \epsilon_{ij}^s \quad (\text{TRIM})$$

where $\epsilon_{ij}^s \doteq \frac{\tilde{\epsilon}_{ij}}{\|\bar{\mathbf{a}}_{ij}\|}$, and $o_{ij}^s \doteq \frac{\tilde{o}_{ij}}{\|\bar{\mathbf{a}}_{ij}\|}$. It is easy to see that $|\epsilon_{ij}^s| \leq \delta_{ij}/\|\bar{\mathbf{a}}_{ij}\|$ since $|\tilde{\epsilon}_{ij}| \leq \delta_{ij}$. We define $\alpha_{ij} \doteq \delta_{ij}/\|\bar{\mathbf{a}}_{ij}\|$.

Eq. (TRIM) describes a TRIM whose generative model is only function of the unknown scale s .

Remark 5 (Novelty of Invariant Measurements): Ideas similar to the *translation invariant measurements* (TIMs) have been used in recent work [19], [29], [99], [100], [101] while 1) the novel graph-theoretic interpretation of Theorem 4 generalizes previously proposed methods and allows pruning outliers as described in Section VI-D, and 2) the notion of TRIMs is completely new. We also remark that while related work uses invariant measurements to filter-out outliers [19] or to speed up

BnB, [29], [99], we show that they also allow decoupling the estimation of scale, rotation, and translation.

A summary table of the invariant measurements and the corresponding noise bounds is given in Appendix E in [27].

VI. TRUNCATED LEAST SQUARES ESTIMATION AND SEMIDEFINITE RELAXATION (TEASER): OVERVIEW

We propose a decoupled approach to solve in cascade for the scale, the rotation, and the translation in (6). The approach, named *Truncated least squares Estimation And SEMidefinite Relaxation* (TEASER), works as follows:

- 1) We use the TRIMs to estimate the scale \hat{s} .
- 2) We use \hat{s} and the TIMs to estimate the rotation $\hat{\mathbf{R}}$.
- 3) We use \hat{s} and $\hat{\mathbf{R}}$ to estimate the translation $\hat{\mathbf{t}}$ from $(\mathbf{a}_i, \mathbf{b}_i)$ in the original TLS problem (6).

We state each subproblem in the following subsections.

A. Robust Scale Estimation

The generative model (TRIM) describes linear *scalar* measurements s_{ij} of the unknown scale s , affected by bounded noise $|\epsilon_{ij}^s| \leq \alpha_{ij}$ including potential outliers (when $\mathbf{o}_{ij}^s \neq \mathbf{0}$). Again, we estimate the scale given the measurements s_{ij} and the bounds α_{ij} using a TLS estimator:

$$\hat{s} = \arg \min_s \sum_{k=1}^K \min \left(\frac{(s - s_k)^2}{\alpha_k^2}, \bar{c}^2 \right) \quad (11)$$

where for simplicity, we numbered the invariant measurements from 1 to $K = |\mathcal{E}|$ and adopted the notation s_k instead of s_{ij} . Section VII shows that (11) can be solved exactly and in polynomial time via *adaptive voting* (see Algorithm 2).

将其转化到多项式时间进行求解

B. Robust Rotation Estimation

Given the scale estimate \hat{s} produced by the scale estimation, the generative model (TIM) describes measurements $\bar{\mathbf{b}}_{ij}$ affected by bounded noise $|\epsilon_{ij}^s| \leq \delta_{ij}$ including potential outliers (when $\mathbf{o}_{ij}^s \neq \mathbf{0}$). Again, we compute \mathbf{R} from the estimated scale \hat{s} , the TIM measurements $(\bar{\mathbf{a}}_{ij}, \bar{\mathbf{b}}_{ij})$ and the bounds δ_{ij} using a TLS estimator:

$$\hat{\mathbf{R}} = \arg \min_{\mathbf{R} \in \text{SO}(3)} \sum_{k=1}^K \min \left(\frac{\|\bar{\mathbf{b}}_k - \hat{s}\mathbf{R}\bar{\mathbf{a}}_k\|^2}{\delta_k^2}, \bar{c}^2 \right) \quad (12)$$

where for simplicity, we numbered the measurements from 1 to $K = |\mathcal{E}|$ and adopted the notation $\bar{\mathbf{a}}_k, \bar{\mathbf{b}}_k$ instead of $\bar{\mathbf{a}}_{ij}, \bar{\mathbf{b}}_{ij}$. Problem (12) is known as the *Robust Wahba* or *Robust Rotation Search* problem [40]. Section VIII shows that (12) can be solved exactly and in polynomial time (in practical problems) *via a tight semidefinite relaxation*.

C. Robust Component-Wise Translation Estimation

After obtaining the scale and rotation estimates \hat{s} and $\hat{\mathbf{R}}$ by solving (11)–(12), we can substitute them back into problem (6) to estimate the translation \mathbf{t} . Although (6) operates on the ℓ_2 norm of the vector, we propose to solve for the translation component-wise, i.e., we compute the entries t_1, t_2, t_3 of \mathbf{t}

independently:

$$\hat{t}_j = \arg \min_{t_j} \sum_{i=1}^N \min \left(\frac{(t_j - [\mathbf{b}_i - \hat{s}\hat{\mathbf{R}}\mathbf{a}_i]_j)^2}{\beta_i^2}, \bar{c}^2 \right) \quad (13)$$

感觉到这里，这篇文章大概思路是有了，接下来是一些优化的trick了

for $j = 1, 2, 3$, and where $[\cdot]_j$ denotes the j th entry of a vector. Since $\mathbf{b}_i - \hat{s}\hat{\mathbf{R}}\mathbf{a}_i$ is a known vector in this stage, it is easy to see that (13) is a scalar TLS problem. Therefore, similarly to (11), Section VII shows that (13) can be solved exactly and in polynomial time via *adaptive voting* (see Algorithm 2). The interested reader can find a discussion on component-wise versus full TLS translation estimation in Appendix Q in [27].

D. Boosting Performance: Max Clique Inlier Selection (MCIS)

While in principle, we could simply execute the cascade of scale, rotation, and translation estimation described above, our graph-theoretic interpretation of Theorem 4 affords further opportunities to prune outliers.

Consider the TRIMs as edges in the complete graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ (where the vertices \mathcal{V} are the correspondences and the edge set \mathcal{E} induces the TIMs and TRIMs per Theorem 4). After estimating the scale (11) (discussed in Section VII), we can prune the edges (i, j) in the graph whose associated TRIM s_{ij} have been classified as outliers by the TLS formulation (*i.e.*, $|s_{ij} - \hat{s}| > \bar{c}\alpha_{ij}$). This allows us to obtain a pruned graph $\mathcal{G}'(\mathcal{V}, \mathcal{E}')$, with $\mathcal{E}' \subseteq \mathcal{E}$, where gross outliers are discarded. The following result ensures that inliers form a clique in the graph $\mathcal{G}'(\mathcal{V}, \mathcal{E}')$, enabling an even more substantial rejection of outliers.

Theorem 6 (MMCIS): Edges corresponding to inlier TIMs form a clique in \mathcal{E}' , and there is at least one maximal clique in \mathcal{E}' that contains all the inliers.

A proof of Theorem 6 is presented in Appendix F in [27]. Theorem 6 allows us to prune outliers by finding the maximal cliques of $\mathcal{G}'(\mathcal{V}, \mathcal{E}')$. Similar idea has been explored for rigid body motion segmentation [102]. Although finding the maximal cliques of a graph takes exponential time in general, there exist efficient approximation algorithms that scale to graphs with millions of nodes [103]–[105]. Under high outlier rates, the graph $\mathcal{G}'(\mathcal{V}, \mathcal{E}')$ is sparse and the maximal clique problem can be solved quickly in practice [106]. Therefore, in this article, after performing scale estimation and removing the corresponding gross outliers, we compute the maximal clique with largest cardinality, *i.e.*, the *maximum clique*, as the inlier set to pass to rotation estimation. Section XI-A shows that this method drastically reduces the number of outliers.

E. Pseudocode of TEASER

The pseudocode of TEASER is summarized in Algorithm 1.

The following Sections VII and VIII describe how to implement the functions in lines 7, 9, 10 of Algorithm 1. In particular, we show how to obtain global and robust estimates of scale (estimate_s) and translation (estimate_t) in Section VII, and rotation (estimate_R) in Section VIII.

Algorithm 1: Truncated Least Squares Estimation And SEMidefinite Relaxation (TEASER).

- 1: **Input:** points $(\mathbf{a}_i, \mathbf{b}_i)$ and bounds β_i ($i = 1, \dots, N$), threshold \bar{c}^2 (default: $\bar{c}^2 = 1$), graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ (default: \mathcal{G} describes the complete graph);
 - 2: **Output:** $\hat{s}, \hat{\mathbf{R}}, \hat{\mathbf{t}}$;
 - 3: % Compute TIM and TRIM
 - 4: $\bar{\mathbf{b}}_{ij} = \mathbf{b}_j - \mathbf{b}_i$, $\bar{\mathbf{a}}_{ij} = \mathbf{a}_j - \mathbf{a}_i$, $\delta_{ij} = \beta_i + \beta_j$, $\forall (i, j) \in \mathcal{E}$
 - 5: $s_{ij} = \frac{\|\bar{\mathbf{b}}_{ij}\|}{\|\bar{\mathbf{a}}_{ij}\|}$, $\alpha_{ij} = \frac{\delta_{ij}}{\|\bar{\mathbf{a}}_{ij}\|}$, $\forall (i, j) \in \mathcal{E}$
 - 6: % Decoupled estimation of $s, \mathbf{R}, \mathbf{t}$
 - 7: $\hat{s} = \text{estimate_s}(\{s_{ij}, \alpha_{ij} : \forall (i, j) \in \mathcal{E}\}, \bar{c}^2)$
 - 8: $\mathcal{G}'(\mathcal{V}', \mathcal{E}') = \text{maxClique}(\mathcal{G}(\mathcal{V}, \mathcal{E}'))$ % prune outliers
 - 9: $\hat{\mathbf{R}} = \text{estimate_R}(\{\bar{\mathbf{a}}_{ij}, \bar{\mathbf{b}}_{ij}, \delta_{ij} : \forall (i, j) \in \mathcal{E}'\}, \bar{c}^2, \hat{s})$
 - 10: $\hat{\mathbf{t}} = \text{estimate_t}(\{\mathbf{a}_i, \mathbf{b}_i, \beta_i : i \in \mathcal{V}'\}, \bar{c}^2, \hat{s}, \hat{\mathbf{R}})$
 - 11: **return:** $\hat{s}, \hat{\mathbf{R}}, \hat{\mathbf{t}}$
-

VII. ROBUST SCALE AND TRANSLATION ESTIMATION: ADAPTIVE VOTING

In this section, we propose an *adaptive voting* algorithm to solve *exactly* the robust scale estimation and the robust component-wise translation estimation.

A. Adaptive Voting for Scalar TLS Estimation

Both the scale estimation (11) and the component-wise translation estimation (13) resort to finding a TLS estimate of an unknown *scalar* given a set of outlier-corrupted measurements. Using the notation for scale estimation (11), the following theorem shows that one can solve *scalar* TLS *estimation* in polynomial time by a simple enumeration.

Theorem 7 (Optimal Scalar TLS Estimation): Consider the scalar TLS problem in (11). For a given $s \in \mathbb{R}$, define the *consensus set of s* as $\mathcal{I}(s) = \{k : \frac{(s-s_k)^2}{\alpha_k^2} \leq \bar{c}^2\}$. Then, for any $s \in \mathbb{R}$, there are at most $2K - 1$ different nonempty consensus sets. If we name these sets $\mathcal{I}_1, \dots, \mathcal{I}_{2K-1}$, then the solution of (11) can be computed by enumeration as

$$\hat{s} = \arg \min \left\{ f_s(\hat{s}_i) : \hat{s}_i = \left(\sum_{k \in \mathcal{I}_i} \frac{1}{\alpha_k^2} \right)^{-1} \sum_{k \in \mathcal{I}_i} \frac{s_k}{\alpha_k^2}, \forall i \right\} \quad (14)$$

where $f_s(\cdot)$ is the objective function of (11).

Theorem 7, whose proof is given in Appendix G in [27], is based on the insight that the consensus set can only change at the boundaries of the intervals $[s_k - \alpha_k \bar{c}, s_k + \alpha_k \bar{c}]$ [see Fig. 3(a)] and there are at most $2K$ such boundaries. The theorem also suggests a straightforward *adaptive voting* algorithm to solve (11), with pseudocode given in Algorithm 2. The algorithm first builds the boundaries of the intervals shown in Fig. 3(a) (line 4). Then, for each interval, it evaluates the consensus set [line 12, see also Fig. 3(b)]. Since the consensus set does not change within an interval, we compute it at the interval centers [line 6, see also Fig. 3(b)]. Finally, the cost of each consensus set is computed and the smallest cost is returned as optimal solution (line 17).

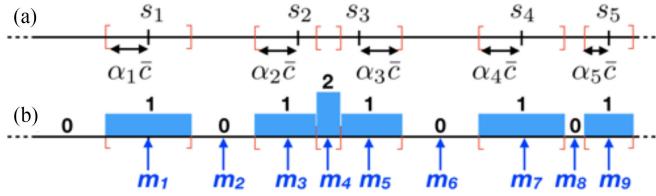


Fig. 3. (a) confidence interval for each measurement s_k (every s in the k th interval satisfies $\frac{(s-s_k)^2}{\alpha_k^2} \leq \bar{c}^2$; (b) cardinality of the consensus set for every s and middle-points m_i for each interval with constant consensus set.

Algorithm 2: Adaptive Voting.

```

1 Input:  $s_k, \alpha_k, \bar{c}$ ;
2 Output:  $\hat{s}$ , scale estimate solving (11);
3 % Define boundaries and sort
4  $v = \text{sort}([s_1 - \alpha_1 \bar{c}, s_1 + \alpha_1 \bar{c}, \dots, s_K - \alpha_K \bar{c}, s_K + \alpha_K \bar{c}])$ 
5 % Compute middle points
6  $m_i = \frac{v_i + v_{i+1}}{2}$  for  $i = 1, \dots, 2K - 1$ 
7 % Voting
8 for  $i = 1, \dots, 2K - 1$  do
9    $\mathcal{I}_i = \emptyset$ 
10  for  $k = 1, \dots, K$  do
11    if  $m_i \in [s_k - \alpha_k \bar{c}, s_k + \alpha_k \bar{c}]$  then
12       $\mathcal{I}_i = \mathcal{I}_i \cup \{k\}$  % add to consensus set
13    end
14  end
15 end
16 % Enumerate consensus sets and return best
17 return:  $\hat{s}$  from Eq. (14).
  
```

Remark 8 (Adaptive Voting): The adaptive voting algorithm generalizes the histogram voting method of Scaramuzza [107] to adaptively adjust the bin size in order to obtain an optimal solution and 2) to solve a TLS (rather than Consensus Maximization) formulation. Adaptive voting can be also used for Consensus Maximization, by simply returning the largest consensus set \mathcal{I}_i in Algorithm 2. We also refer the interested reader to the paper of Liu and Jiang [95], who recently developed a similar algorithm in an independent effort and provide a generalization to 2D TLS estimation problems.

In summary, the function `estimate_s` in Algorithm 1 calls Algorithm 2 to compute the optimal scale estimate \hat{s} , and the function `estimate_t` in Algorithm 1 calls Algorithm 2 three times (one for each entry of t) and returns the translation estimate $\hat{t} = [\hat{t}_1 \ \hat{t}_2 \ \hat{t}_3]^T$.

VIII. ROBUST ROTATION ESTIMATION: SEMIDEFINITE RELAXATION AND FAST CERTIFICATES

This section describes how to compute an optimal solution to problem (12) or certify that a given rotation estimate is globally optimal. TLS estimation is NP-hard according to Remark 3 and the robust rotation estimation (12) has the additional complexity of involving a nonconvex domain [*i.e.*, $\text{SO}(3)$]. This section provides a surprising result: we can compute globally optimal solutions to (3) (or certify that a given estimate is optimal) in polynomial time in virtually all practical problems using a *tight*

semidefinite (SDP) relaxation. In other words, while the NP-hardness implies the presence of worst-case instances that are not solvable in polynomial time, these instances are not observed to frequently occur in practice.

We achieve this goal in three steps. First, we show that the robust rotation estimation problem (12) can be reformulated as a *Quadratically Constrained Quadratic Program* (QCQP) by adopting a *quaternion* formulation and using a technique we name *binary cloning* (see Section VIII-A). Second, we show how to obtain a *semidefinite relaxation* and relax the nonconvex QCQP into a convex SDP with *redundant constraints* (see Section VIII-B). The SDP relaxation enables solving (12) with global optimality guarantees in polynomial time. Third, we show that the relaxation also enables a fast *optimality certifier*, which, given a rotation guess, can test if the rotation is the optimal solution to (12) (see Section VIII-C). The latter will be instrumental in developing a fast and certifiable registration approach that circumvents the time-consuming task of solving a large relaxation with existing SDP solvers. The derivation in Sections VIII-A and VIII-B is borrowed from our previous work [40], while the fast certification in Section VIII-C has not been presented before.

A. Robust Rotation Estimation as a QCQP

This section rewrites problem (3) as a QCQP. Using the quaternion preliminaries introduced in Section III—and in particular (3)—it is easy to rewrite (12) using unit quaternions:

$$\min_{\mathbf{q} \in \mathbb{S}^3} \sum_{k=1}^K \min \left(\frac{\|\hat{\mathbf{b}}_k - \mathbf{q} \circ \hat{\mathbf{a}}_k \circ \mathbf{q}^{-1}\|^2}{\delta_k^2}, \bar{c}^2 \right) \quad (15)$$

where we defined $\hat{\mathbf{a}}_k \doteq [\hat{s} \bar{a}_k^T 0]^T$ and $\hat{\mathbf{b}}_k \doteq [\bar{b}_k^T 0]^T$, and “ \circ ” denotes the quaternion product. The main advantage of using (15) is that we replaced the set $\text{SO}(3)$ with a simpler set, the four-dimensional unit sphere \mathbb{S}^3 .

From TLS to Mixed-Integer Programming: Problem (15) is hard to solve *globally*, due to the nonconvexity of both the cost function and the domain \mathbb{S}^3 . As a first step towards obtaining a QCQP, we expose the nonconvexity of the cost by rewriting the TLS cost using *binary variables*. In particular, we rewrite the inner “min” in (15) using the following property, that holds for any pair of scalars x and y :

$$\min(x, y) = \min_{\theta \in \{+1, -1\}} \frac{1+\theta}{2}x + \frac{1-\theta}{2}y. \quad (16)$$

Equation (16) can be verified to be true by inspection: the right-hand-side returns x (with minimizer $\theta = +1$) if $x < y$, and y (with minimizer $\theta = -1$) if $x > y$. This enables us to rewrite problem (15) as a mixed-integer program, including the quaternion \mathbf{q} and binary variables θ_k , $i = 1, \dots, K$

$$\min_{\substack{\mathbf{q} \in \mathbb{S}^3 \\ \theta_k = \{\pm 1\}}} \sum_{k=1}^K \frac{1+\theta_k}{2} \frac{\|\hat{\mathbf{b}}_k - \mathbf{q} \circ \hat{\mathbf{a}}_k \circ \mathbf{q}^{-1}\|^2}{\delta_k^2} + \frac{1-\theta_k}{2} \bar{c}^2. \quad (17)$$

The reformulation is related to the *Black–Rangarajan duality* between robust estimation and line processes [54]: the TLS cost is an extreme case of robust function that results in a binary line process. Intuitively, the binary variables $\{\theta_k\}_{k=1}^N$ in problem (17) decide whether a given measurement k is an inlier ($\theta_k = +1$) or an outlier ($\theta_k = -1$). Exposing the nonconvexity *如果*是outlier, 则是取小的部分, 也就是 c^2

of TLS cost function using binary variables makes the TLS cost more favorable than other nonconvex robust costs (e.g., the Geman–McClure cost used in FGR [56]), because it enables binary cloning (Proposition 9) and semidefinite relaxation (see Section VIII-B).

From Mixed-Integer to Quaternions: Now we convert the mixed-integer program (17) to an optimization over $K + 1$ quaternions. The intuition is that, if we define extra quaternions $\mathbf{q}_k \doteq \theta_k \mathbf{q}$, we can rewrite (17) as a function of \mathbf{q} and \mathbf{q}_k ($k = 1, \dots, K$). This is a key step towards getting a quadratic cost (Proposition 10) and is formalized as follows.

Proposition 9 (Binary Cloning): The mixed-integer program (17) is equivalent (in the sense that they admit the same optimal solution \mathbf{q}) to the following optimization:

$$\begin{aligned} \min_{\substack{\mathbf{q} \in \mathbb{S}^3 \\ \mathbf{q}_k = \{\pm \mathbf{q}\}}} \quad & \sum_{k=1}^K \frac{\|\hat{\mathbf{b}}_k - \mathbf{q} \circ \hat{\mathbf{a}}_k \circ \mathbf{q}^{-1} + \mathbf{q}^\top \mathbf{q}_k \hat{\mathbf{b}}_k - \mathbf{q} \circ \hat{\mathbf{a}}_k \circ \mathbf{q}_k^{-1}\|^2}{4\delta_k^2} \\ & + \frac{1 - \mathbf{q}^\top \mathbf{q}_k}{2} \bar{c}^2. \end{aligned} \quad (18)$$

which involves $K + 1$ quaternions (\mathbf{q} and \mathbf{q}_k , $i = 1, \dots, K$).

While a formal proof is given in Appendix H in [27], it is fairly easy to see that if $\mathbf{q}_k = \{\pm \mathbf{q}\}$, or equivalently, $\mathbf{q}_k = \theta_k \mathbf{q}$ with $\theta_k \in \{\pm 1\}$, then $\mathbf{q}_k^\top \mathbf{q} = \theta_k$, and $\mathbf{q} \circ \hat{\mathbf{a}}_k \circ \mathbf{q}_k^{-1} = \theta_k (\mathbf{q} \circ \hat{\mathbf{a}}_k \circ \mathbf{q}^{-1})$, which exposes the relation between (17) and (18). We dubbed the reparametrization (18) *binary cloning* since now we created a “clone” \mathbf{q}_k for each measurement, such that $\mathbf{q}_k = \mathbf{q}$ for inliers (recall $\mathbf{q}_k = \theta_k \mathbf{q}$) and $\mathbf{q}_k = -\mathbf{q}$ for outliers.

From Quaternions to QCQP: We conclude this section by showing that (18) can be actually written as a QCQP. This observation is non-trivial since (18) has a *quartic* cost and $\mathbf{q}_k = \{\pm \mathbf{q}\}$ is not in the form of a quadratic constraint. The reformulation as a QCQP is given in the following.

Proposition 10 (Binary Cloning as a QCQP): Define a single column vector $\mathbf{x} = [\mathbf{q}^\top \mathbf{q}_1^\top \dots \mathbf{q}_K^\top]^\top$ stacking all variables in Problem (18). Then, Problem (18) is equivalent (in the sense that they admit the same optimal solution \mathbf{q}) to the following QCQP:

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^{4(K+1)}} \quad & \mathbf{x}^\top \mathbf{Q} \mathbf{x} \\ \text{s.t.} \quad & \mathbf{x}_q^\top \mathbf{x}_q = 1 \\ & \mathbf{x}_{q_k}^\top \mathbf{x}_{q_k} = \mathbf{x}_q^\top \mathbf{x}_q, \forall k = 1, \dots, K. \end{aligned} \quad (19)$$

where $\mathbf{Q} \in \mathcal{S}^{4(K+1)}$ is a known symmetric matrix that depends on the TIM measurements $\bar{\mathbf{a}}_k$ and $\bar{\mathbf{b}}_k$ (the explicit expression is given in Appendix I in [27]), and the notation \mathbf{x}_q (resp. \mathbf{x}_{q_k}) denotes the 4-D subvector of \mathbf{x} corresponding to \mathbf{q} (resp. \mathbf{q}_k).

A complete proof of Proposition 10 is given in Appendix I in [27]. Intuitively, 1) we developed the squares in the cost function (18), 2) we used the properties of unit quaternions (see Section III) to simplify the expression to a quadratic cost, and 3) we adopted the more compact notation afforded by the vector \mathbf{x} to obtain (19).

B. Semidefinite Relaxation

Problem (19) writes the TLS rotation estimation problem (12) as a QCQP. Problem (19) is still a non-convex problem (quadratic

equality constraints are nonconvex). Here, we develop a *tight* convex SDP relaxation for problem (19).

The crux of the relaxation consists in rewriting problem (19) as a function of the following matrix:

$$\mathbf{Z} = \mathbf{x} \mathbf{x}^\top = \begin{bmatrix} \mathbf{q} \mathbf{q}^\top & \mathbf{q} \mathbf{q}_1^\top & \cdots & \mathbf{q} \mathbf{q}_K^\top \\ * & \mathbf{q}_1 \mathbf{q}_1^\top & \cdots & \mathbf{q}_1 \mathbf{q}_K^\top \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & \mathbf{q}_K \mathbf{q}_K^\top \end{bmatrix} \in \mathcal{S}_+^{4(K+1)}. \quad (20)$$

For this purpose, we note that the objective function of (19) is a linear function of \mathbf{Z}

$$\mathbf{x}^\top \mathbf{Q} \mathbf{x} = \text{tr}(\mathbf{Q} \mathbf{x} \mathbf{x}^\top) = \text{tr}(\mathbf{Q} \mathbf{Z}) \quad (21)$$

and that $\mathbf{x}_q \mathbf{x}_q^\top = [\mathbf{Z}]_{00}$, where $[\mathbf{Z}]_{00}$ denotes the top-left 4×4 diagonal block of \mathbf{Z} , and $\mathbf{x}_{q_k} \mathbf{x}_{q_k}^\top = [\mathbf{Z}]_{kk}$, where $[\mathbf{Z}]_{kk}$ denotes the k th diagonal block of \mathbf{Z} (we number the row and column block index from 0 to K for notation convenience). Since any matrix in the form $\mathbf{Z} = \mathbf{x} \mathbf{x}^\top$ is a PSD rank-1 matrix, we obtain:

Proposition 11 (Matrix Formulation of Binary Cloning): Problem (19) is equivalent (in the sense that optimal solutions of a problem can be mapped to optimal solutions of the other) to the following nonconvex rank-constrained program:

$$\begin{aligned} \min_{\mathbf{Z} \succeq 0} \quad & \text{tr}(\mathbf{Q} \mathbf{Z}) \\ \text{s.t.} \quad & \text{tr}([\mathbf{Z}]_{00}) = 1 \\ & [\mathbf{Z}]_{kk} = [\mathbf{Z}]_{00}, \forall k = 1, \dots, K \\ & \text{rank}(\mathbf{Z}) = 1. \end{aligned} \quad (22)$$

The proof is given in Appendix J in [27]. At this point, we are ready to develop our SDP relaxation by dropping the nonconvex rank-1 constraint and adding redundant constraints.

Proposition 12 (SDP Relaxation With Redundant Constraints): The following SDP is a convex relaxation of (22)

$$\begin{aligned} \min_{\mathbf{Z} \succeq 0} \quad & \text{tr}(\mathbf{Q} \mathbf{Z}) \\ \text{s.t.} \quad & \text{tr}([\mathbf{Z}]_{00}) = 1 \\ & [\mathbf{Z}]_{kk} = [\mathbf{Z}]_{00}, \forall k = 1, \dots, K \\ & [\mathbf{Z}]_{ij} = [\mathbf{Z}]_{ij}^\top, \forall 0 \leq i < j \leq K. \end{aligned} \quad (23)$$

The redundant constraints in the last row of (23) enforce all the off-diagonal 4×4 blocks of \mathbf{Z} to be symmetric, which is a redundant constraint for (19), since

$$[\mathbf{Z}]_{ij} = \mathbf{q}_i \mathbf{q}_j^\top = (\theta_i \mathbf{q})(\theta_j \mathbf{q})^\top = \theta_i \theta_j \mathbf{q} \mathbf{q}^\top \quad (24)$$

is indeed a signed copy of $\mathbf{q} \mathbf{q}^\top$ and must be symmetric. Although being redundant for the QCQP (19), these constraints are critical in tightening the SDP relaxation (cf. [40]). Appendix K in [27] provides a different approach to develop the convex relaxation using Lagrangian duality theory.

The following theorem provides readily checkable *a posteriori* conditions under which (23) computes an optimal solution to robust rotation estimation.

Theorem 13 (Global Optimality Guarantee): Let \mathbf{Z}^* be the optimal solution of (23). If $\text{rank}(\mathbf{Z}^*) = 1$, then the SDP relaxation is said to be *tight*, and:

- 1) the optimal cost of (23) matches the optimal cost of the QCQP (19),

- 2) Z^* can be written as $Z^* = (\mathbf{x}^*)^T(\mathbf{x}^*)$, where $\mathbf{x}^* = [(\mathbf{q}^*)^T, (\mathbf{q}_1^*)^T, \dots, (\mathbf{q}_K^*)^T]^T$, with $\mathbf{q}_k^* = \theta_k^* \mathbf{q}^*$, $\theta_k^* \in \{\pm 1\}$,
 3) $\pm \mathbf{x}^*$ are the two unique global minimizers of the original QCQP (19).

A formal proof is given in Appendix K in [27]. Theorem 13 states that if the SDP (23) admits a rank-1 solution, then one can extract the unique global minimizer (remember: \mathbf{q} and $-\mathbf{q}$ represent the same rotation) of the nonconvex QCQP (19) from the rank-1 decomposition of the SDP solution.

In practice, we observe that (23) always returns a rank-1 solution: in Section XI, we empirically demonstrate the tightness of (23) in the face of noise and extreme outlier rates ($> 95\%$ [40]), and show its superior performance compared to a similar SDP relaxation using a rotation matrix parametrization [see [39, eq. (13) in Sec. V.B]. As it is often the case, we can only partially justify with theoretical results the phenomenal performance observed in practice. The interested reader can find a formal proof of tightness in the low-noise and outlier-free case in [40, Th. 7 and Proposition 8].

C. Fast Global Optimality Certification

Despite the superior robustness and (*a posteriori*) global optimality guarantees, the approach outlined in the previous section requires solving the large-scale SDP (23), which is known to be computationally expensive (e.g., solving the SDP (23) for $K = 100$ takes about 1200 s with MOSEK [108]). On the other hand, there exist fast heuristics that compute high-quality solutions to the nonconvex TLS rotation problem (12) in milliseconds. For example, the GNC approach in [35] computes globally optimal TLS solutions with high probability when the outlier ratio is below 80%, while not being certifiable.

Therefore, in this section, we ask the question: *given an estimate from a fast heuristic, such as GNC², can we certify the optimality of the estimate or reject it as suboptimal?* In other words: can we make GNC certifiable? The answer is yes: we can leverage the insights of Section VIII-B to obtain a fast optimality certifier that is orders of magnitude faster than solving the relaxation (23).³ The pseudocode is presented in Algorithm 3 and the following theorem proves its soundness.

Theorem 14 (Optimality Certification): Given a feasible (but not necessarily optimal) solution $(\hat{\mathbf{q}}, \hat{\theta}_1, \dots, \hat{\theta}_K)$ of the TLS rotation estimation problem (15), denote with $\hat{\mathbf{x}} = [\hat{\mathbf{q}}^T, \hat{\theta}_1 \hat{\mathbf{q}}^T, \dots, \hat{\theta}_K \hat{\mathbf{q}}^T]^T$, and $\hat{\mu} = \hat{\mathbf{x}}^T Q \hat{\mathbf{x}}$, the corresponding solution and cost of the QCQP (19). Then, Algorithm 3 produces a *suboptimality bound* η that satisfies

$$\frac{\hat{\mu} - \mu^*}{\hat{\mu}} \leq \eta \quad (25)$$

where μ^* is the (unknown) global minimum of the nonconvex TLS rotation estimation problem (15) and the QCQP (19). Moreover, when the relaxation (23) is tight and the parameter γ in line 3 satisfies $0 < \gamma < 2$, and if $\hat{\mathbf{q}} = \mathbf{q}^*$, $\hat{\theta}_k = \theta_k^*$ is a global minimizer of the TLS problem (15), then the suboptimality bound $\eta^{(t)}$ (line 3) converges to zero.

²Our certification approach also applies to RANSAC, which however has the downside of being non-deterministic and typically less accurate than GNC.

³GNC (or RANSAC) can be seen as a fast *primal* solver that returns a rank-1 guess of (23) by estimating \mathbf{x} (i.e., \mathbf{q} and θ_k , $k = 1, \dots, K$) and $Z = \mathbf{x}\mathbf{x}^T$, and the optimality certifier can be seen as a fast *dual-from-primal* solver that attempts to compute a dual optimality certificate for the primal rank-1 guess.

Algorithm 3: Optimality Certification.

```

1 Input: a feasible solution  $(\hat{\mathbf{q}}, \hat{\theta}_1, \dots, \hat{\theta}_K)$  attaining
   cost  $\hat{\mu}$  in the TLS rotation estimation problem (15);
   homogeneous and scaled TIMs  $(\hat{\mathbf{a}}_k, \hat{\mathbf{b}}_k)$ , noise bounds
    $\delta_k$  and  $\bar{c}^2$ ; maximum number of iterations  $T$  (default
    $T = 200$ ); desired relative suboptimality gap  $\bar{\eta}$ 
   (default  $\bar{\eta} = 0.1\%$ ); initial suboptimality gap
    $\eta = +\infty$ ; relaxation parameter  $\gamma$  ( $0 < \gamma < 2$ )
2 Output: bound on relative suboptimality gap  $\eta$ 
3 % Compute data matrix  $\mathbf{Q}$  (Proposition 10)
4  $\mathbf{Q} = \text{get\_Q}(\hat{\mathbf{a}}_k, \hat{\mathbf{b}}_k, \delta_k, \bar{c}^2)$ 
5 % Compute  $\bar{\mathbf{x}}$ ,  $\bar{\mathbf{Q}}$  (Appendix L)
6  $\bar{\mathbf{x}} = [1, \hat{\theta}_1, \dots, \hat{\theta}_K]^T \otimes [0 \ 0 \ 0 \ 1]^T$ ,  $\bar{\mathbf{Q}} = \hat{\Omega}_q^T \mathbf{Q} \hat{\Omega}_q$ 
7 % Compute initial dual certificate  $\bar{\mathbf{M}}^{(0)} \in \bar{\mathcal{L}}$ 
8  $\bar{\mathbf{M}}^{(0)} = \bar{\mathbf{Q}} - \hat{\mu} \mathbf{J} + \text{get\_Delta0}(\hat{\mathbf{a}}_k, \hat{\mathbf{b}}_k, \delta_k, \bar{c}^2, \bar{\mathbf{Q}}, \bar{\mathbf{x}})$ 
9 % Douglas-Rachford Splitting
10 for  $t = 0, \dots, T$  do
11    $\bar{\mathbf{M}}_{\mathcal{S}_+}^{(t)} = \Pi_{\mathcal{S}_+}(\bar{\mathbf{M}}^{(t)})$  % Project to  $\mathcal{S}_+$ 
12    $\bar{\mathbf{M}}_{\bar{\mathcal{L}}}^{(t)} = \Pi_{\bar{\mathcal{L}}}(2\bar{\mathbf{M}}_{\mathcal{S}_+}^{(t)} - \bar{\mathbf{M}}^{(t)})$  % Project to  $\bar{\mathcal{L}}$ 
13    $\bar{\mathbf{M}}^{(t+1)} = \bar{\mathbf{M}}^{(t)} + \gamma(\bar{\mathbf{M}}_{\bar{\mathcal{L}}}^{(t)} - \bar{\mathbf{M}}_{\mathcal{S}_+}^{(t)})$  % Relaxation
14   % Compute relative suboptimality bound
15    $\lambda_1^{(t)} = \text{get\_min\_eig}(\bar{\mathbf{M}}_{\bar{\mathcal{L}}}^{(t)})$ ,  $\eta^{(t)} = \frac{|\lambda_1^{(t)}|(K+1)}{\hat{\mu}}$ 
16   if  $\eta^{(t)} < \eta$  then  $\eta = \eta^{(t)}$  end
17   if  $\eta < \bar{\eta}$  then break end
18 end
19 return:  $\eta$ 

```

Although a complete derivation of Algorithm 3 and the proof of Theorem 14 is postponed to Appendix L in [27], now we describe our key intuitions. The first insight is that, given a globally optimal solution $(\mathbf{q}^*, \theta_1^*, \dots, \theta_K^*)$ to the nonconvex TLS problem (15), whenever the SDP relaxation (23) is tight, $Z^* = (\mathbf{x}^*)^T(\mathbf{x}^*)$, where $\mathbf{x}^* = [(\mathbf{q}^*)^T, (\theta_1^* \mathbf{q}^*)^T, \dots, (\theta_K^* \mathbf{q}^*)^T]$, is a globally optimal solution to the convex SDP (23). Therefore, we can compute a certificate of global optimality from the Lagrangian dual SDP of the primal SDP (23) [109, Sec. 5].⁴

The second insight is that Lagrangian duality shows that the dual certificate is a matrix at the intersection between the PSD cone \mathcal{S}_+ and an affine subspace $\bar{\mathcal{L}}$ (whose expression is derived in Appendix L in [27]). Algorithm 3 looks for such a matrix $(\bar{\mathbf{M}}_{\bar{\mathcal{L}}}^{(t)})$ in line (3)) using (DRS) [36], [110]–[114] (line 11–13), starting from a cleverly chosen initial guess (line 3). Since DRS guarantees convergence to a point in $\mathcal{S}_+ \cap \mathcal{M}$ from any initial guess whenever the intersection is nonempty, Algorithm 3 guarantees to certify global optimality if provided with an optimal solution, whenever the relaxation (23) is tight.

The third insight is that even when the provided solution is not globally optimal, or when the relaxation is not tight, Algorithm 3 can still compute a *suboptimality bound* from the *minimum eigenvalue of the dual certificate matrix* (line 15), which is informative of the quality of the candidate solution. To the best of our knowledge, this is the first result that can

⁴The tight SDP relaxation ensures certifying global optimality of the non-convex problem (NP-hard) is equivalent to certifying global optimality of the convex SDP (tractable).

assess the suboptimality of outlier-robust estimation beyond the instances commonly investigated in statistics [34].

Finally, we further speed up computation by exploiting the problem structure. Appendix M in [27] shows that the projections on lines 11 and 12 of Algorithm 3 can be computed in closed form, and Appendix V provides the expression of the initial guess in line 8. Section XI shows the effectiveness of Algorithm 3 in certifying the solutions from GNC [35].

IX. PERFORMANCE GUARANTEES

This section establishes theoretical guarantees for TEASER. While Theorems 7, 13, and 14 establish when we obtain a globally optimal solution from optimization problems involved in TEASER, this section uses these theorems to bound the distance between the estimate produced by TEASER and the ground truth. Section IX-A investigates the noiseless (but outlier-corrupted) case, while Section IX-B considers the general case with noise and outliers.

A. Exact Recovery With Outliers and Noiseless Inliers

We start by analyzing the case in which the inliers are noiseless, since it allows deriving stronger performance guarantees and highlights the fundamental difference the outlier generation mechanism can make. As we show below, in a noiseless case with randomly outliers generated, TEASER is able to recover the ground-truth transformation from only three inliers (and an arbitrary number of outliers!). On the other hand, when the outlier generation is adversarial, TEASER (and any other approach) requires that more than 50% of the measurements are inliers in order to retrieve the ground truth.

Theorem 15 (Estimation Contract With Noiseless Inliers and Random Outliers): Assume (i) the set of correspondences contains at least three noiseless noncollinear and distinct⁵ inliers, where for an inlier i , it holds $\mathbf{b}_i = s^\circ \mathbf{R}^\circ \mathbf{a}_i + \mathbf{t}^\circ$, and $(s^\circ, \mathbf{R}^\circ, \mathbf{t}^\circ)$ denotes the ground-truth transformation we want to recover. Assume (ii) the outliers are in *generic position* (e.g., they are perturbed by random noise). If (iii) the inliers computed by TEASER (in each subproblem) have zero residual error for the corresponding subproblem, and (iv) the rotation subproblem produces a valid certificate (in the sense of Theorems 13 and 14), then with probability 1 the output $(\hat{s}, \hat{\mathbf{R}}, \hat{\mathbf{t}})$ of TEASER exactly recovers the ground truth, i.e., $\hat{s} = s^\circ, \hat{\mathbf{R}} = \mathbf{R}^\circ, \hat{\mathbf{t}} = \mathbf{t}^\circ$.

A proof of Theorem 15 is given in Appendix N in [27]. Conditions (iii) and (iv) can be readily checked using the solution computed by TEASER, in the spirit of certifiable algorithms. Assumptions (i) and (ii) only ensure that the problem is well-posed. In particular assumption (i) is the same assumption required to compute a unique transformation without outliers [14]. Assumption (ii) is trickier since it requires *assuming* the outliers to be “generic” rather than adversarial: intuitively, if we allow an adversary to perturb an arbitrary number of points, s/he can create an identical copy of the points in \mathcal{A} with an arbitrary transformation and induce *any* algorithm into producing an incorrect estimate. Clearly, in such a setup outlier rejection is ill-posed and no algorithm can recover the correct solution.⁶

⁵Distinct in the sense that no two points occupy the same 3-D location.

⁶In such a case, it might still be feasible to enumerate all potential solutions, but the optimal solution of Consensus Maximization or the robust estimator (6) can no longer be guaranteed to be close to the ground truth.

This fundamental limitation is captured by the following theorem, which focuses on the case with adversarial outliers.

Theorem 16 (Estimation Contract with Noiseless Inliers and Adversarial Outliers): Assume (i) the set of correspondences is such that the number of inliers N_{in} and the number of outliers N_{out} satisfy $N_{\text{in}} \geq N_{\text{out}} + 3$ and no three inliers are collinear. If (ii) the inliers computed by TEASER (in each subproblem) have zero residual error for the corresponding subproblem, and (iii) the rotation subproblem produces a valid certificate (in the sense of Theorems 13 and 14), and (iv) TEASER returns a consensus set that satisfies (i), then the output $(\hat{s}, \hat{\mathbf{R}}, \hat{\mathbf{t}})$ of TEASER exactly recovers the ground truth transformation, i.e., $\hat{s} = s^\circ, \hat{\mathbf{R}} = \mathbf{R}^\circ, \hat{\mathbf{t}} = \mathbf{t}^\circ$.

A proof of Theorem 16 is given in Appendix O in [27]. We observe that now Theorem 16 provides an even clearer “contract” for TEASER: as long as the number of inliers is sufficiently larger than the number of outliers, TEASER provides readily-checkable conditions on whether it recovered the ground-truth transformation. If the percentage of inliers is below 50% (i.e., $N_{\text{in}} \leq N_{\text{out}}$), TEASER will attempt to retrieve the transformation that is consistent with the largest set of inliers (in the proof, we show that under the conditions of the theorem, TEASER produces a maximum consensus solution).

To the best of our knowledge, Theorems 15 and 16 provide the first exact recovery results for registration with outliers.

B. Approximate Recovery With Outliers and Noisy Inliers

The case with noisy inliers is strictly more challenging than the noise-free case. Intuitively, the larger the noise, the more blurred is the boundary between inliers and outliers. From the theoretical standpoint, this translates into weaker guarantees.

Theorem 17 (Estimation Contract With Noisy Inliers and Adversarial Outliers): Assume (i) the set of correspondences contains at least four noncoplanar inliers, where for any inlier i , $\|\mathbf{b}_i - s^\circ \mathbf{R}^\circ \mathbf{a}_i - \mathbf{t}^\circ\| \leq \beta$, and $(s^\circ, \mathbf{R}^\circ, \mathbf{t}^\circ)$ denotes the ground-truth transformation. Assume (ii) the inliers belong to the maximum consensus set in each subproblem,⁷ and (iii) the second largest consensus set is “sufficiently smaller” than the maximum consensus set (as formalized in Lemma 23 in [27]). If (iv) the rotation subproblem produces a valid certificate (Theorems 13 and 14), the output $(\hat{s}, \hat{\mathbf{R}}, \hat{\mathbf{t}})$ of TEASER satisfies

$$|\hat{s} - s^\circ| \leq 2 \max_{ij} \alpha_{ij} \quad (26)$$

$$\|\mathbf{s}^\circ \mathbf{R}^\circ - \hat{\mathbf{s}} \hat{\mathbf{R}}\|_F \leq 2\sqrt{3} \frac{\max_{ij} \alpha_{ij}}{\min_{ijk} \sigma_{\min}(\mathbf{U}_{ijk})} \quad (27)$$

$$\|\hat{\mathbf{t}} - \mathbf{t}^\circ\| \leq (9 + 3\sqrt{3})\beta \quad (28)$$

where $\sigma_{\min}(\cdot)$ denotes the smallest singular value of a matrix, and \mathbf{U}_{ijk} is the 3×3 matrix $\mathbf{U}_{ijk} = [\frac{\bar{a}_{ij}}{\|\bar{a}_{ij}\|} \quad \frac{\bar{a}_{ih}}{\|\bar{a}_{ih}\|} \quad \frac{\bar{a}_{ik}}{\|\bar{a}_{ik}\|}]$. Moreover, $\sigma_{\min}(\mathbf{U}_{ijk})$ is nonzero as long as (every 4 points \mathbf{a}_i chosen among) the inliers are not coplanar.

A proof of Theorem 17 is given in Appendix P in [27]. To the best of our knowledge, Theorem (17) provides the first performance bounds for noisy registration problems with outliers. Condition (i) is similar to the standard assumption that guarantees the existence of a unique alignment in the outlier-free case,

⁷In other words: the inliers belong to the largest set \mathcal{S} such that for any $i, j \in \mathcal{S}, |s_{ij} - s| \leq \alpha_{ij}, \|\bar{b}_{ij} - s \bar{R} \bar{a}_{ij}\| \leq \delta_{ij}$, and $\|\mathbf{b}_i - s \mathbf{R} \mathbf{a}_i - \mathbf{t}_l\| \leq \beta_i, \forall l = 1, 2, 3$ for any transformation $(s, \mathbf{R}, \mathbf{t})$.

with the exception that we now require four noncoplanar inliers. While conditions (ii)–(iii) seem different and less intuitive than the ones in Theorem 16, they are a generalization of conditions (i) and (ii) in Theorem 16. Condition (ii) in Theorem 17 ensures that the inliers form a large consensus set (the same role played by condition (i) in Theorem 16). Condition (iii) in Theorem 17 ensures that the largest consensus set cannot be confused with other large consensus sets (the assumption of zero residual played a similar role in Theorem 16). These conditions essentially limit the amount of “damage” that outliers can do by avoiding that they form large sets of mutually consistent measurements.

Remark 18 (Interpretation of the Bounds): The bounds (26)–(28) have a natural geometric interpretation. First of all, we recall from Section V that we defined $\alpha_{ij} = \delta_{ij}/\|\bar{a}_{ij}\|$. Hence α_{ij} can be understood as (the inverse of) a signal to noise ratio: δ_{ij} measures the amount of noise, while $\|\bar{a}_{ij}\| = \|a_j - a_i\|$ measures the size of the point cloud (in terms of distance between two points). Therefore, the scale estimate will worsen when the noise becomes large when compared to the size of the point cloud. The parameter α_{ij} also influences the rotation bound (27), which, however, also depends on the smallest singular value of the matrix $\sigma_{\min}(U_{ijhk})$. The latter measures how far is the point cloud from degenerate configurations. Finally, the translation bound does not depend on the scale of the point cloud, since for a given scale and rotation a single point would suffice to compute a translation (in other words, the distance among points does not play a role there).

The interested reader can find tighter (but more expensive to compute and less intelligible) bounds, under similar assumptions as Theorem 17, in Appendix W in [27].

X. TEASER++: A FAST C++ IMPLEMENTATION

In order to showcase the real-time capabilities of TEASER in real robotics, vision, and graphics applications, we have developed a fast C++ implementation of TEASER, named TEASER++. TEASER++ has been released as an open-source library and can be found at.⁸ This section describes the algorithmic and implementation choices we made in TEASER++.

TEASER++ follows the same decoupled approach described in Section VI. The only exception is that it circumvents the need to solve a large-scale SDP and uses the GNC approach described in [35] for the TLS rotation estimation problem. As described in Section VIII-C, we essentially run GNC and certify *a posteriori* that it retrieved the globally optimal solution using Algorithm 3. This approach is very effective in practice since—as shown in the experimental section—the scale estimation (see Section VI-A) and the max clique inlier selection (see Section VI-D) are already able to remove a large number of outliers, hence GNC only needs to solve a more manageable problem with less outliers (a setup in which it has been shown to be very effective [35]). Therefore, by combining a fast heuristic (GNC) with a scalable optimality certifier (DRS), TEASER++ is the first fast and certifiable registration algorithm in that 1) it enjoys the performance guarantees of Section IX, and 2) it runs in milliseconds in practical problem instances.

TEASER++ uses Eigen3 for fast linear algebra operations. To further optimize for performance, we have implemented a large portion of TEASER++ using shared-memory parallelism. Using

GNC
也可以
认为是一个
模块。
Robust的主要原因是在
scale的阶段
已经对outlier
进行过滤。

⁸[Online]. Available: <https://github.com/MIT-SPARK/TEASER-plusplus>

OpenMP [115], an industry-standard interface for scalable parallel programming, we have parallelized Algorithm 2, as well as the computation of the TIMs and TRIMs. In addition, we use a fast exact parallel maximum clique finder algorithm [116] to find the maximum clique for outlier pruning.

To facilitate rapid prototyping and visualization, in addition to the C++ library, we provide Python and MATLAB bindings. Furthermore, we provide a ROS [117] wrapper to enable easy integration and deployment in real-time robotics applications.

XI. EXPERIMENTS AND APPLICATIONS

The goal of this section is to 1) test the performance of each module presented in this article, including scale, rotation, translation estimation, the MCIS pruning, and the optimality certification Algorithm 3 (see Section XI-A), 2) show that TEASER and TEASER++ outperform the state of the art on standard benchmarks (see Section XI-B), 3) show that TEASER++ also solves SPC problems and dominates existing methods, such as ICP (see Section XI-C), 4) show an application of TEASER for object localization in an RGB-D robotics dataset (see Section XI-D), and 5) show an application of TEASER++ to difficult scan-matching problems with deep-learned correspondences (see Section XI-E).

Implementation Details: TEASER++ is implemented as discussed in Section X. TEASER is implemented in MATLAB, uses cvx [118] to solve the convex relaxation (23), and uses the algorithm in [106] to find the maximum clique in the pruned TIM graph (see Theorem 6). In all tests, we set $\bar{c} = 1$. All tests are run on a laptop with an i7-8850H CPU and 32GB of RAM.

A. Testing TEASER’s and TEASER++’s Modules

Testing Setup: We use the Bunny point cloud from the Stanford 3-D Scanning Repository [98]. The bunny is downsampled to $N = 40$ points (unless specified otherwise) and resized to fit inside the $[0, 1]^3$ cube to create point cloud \mathcal{A} . To create point cloud \mathcal{B} with N correspondences, first a random transformation $(s, \mathbf{R}, \mathbf{t})$ (with $1 \leq s \leq 5$, $\|\mathbf{t}\| \leq 1$, and $\mathbf{R} \in \text{SO}(3)$) is applied according to (4), and then random bounded noise ϵ_i ’s are added (we sample $\epsilon_i \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$ until $\|\epsilon_i\| \leq \beta_i$). We set $\sigma = 0.01$ and $\beta_i = 5.54\sigma$ such that $\mathbb{P}(\|\epsilon_i\|^2 > \beta_i^2) \leq 10^{-6}$ (cf. Remark 20 in Appendix B in [27]). To generate outlier correspondences, we replace a fraction of the b_i ’s with vectors uniformly sampled inside the sphere of radius 5. We test increasing outlier rates from 0% (all inliers) to 90%; we test up to 99% outliers when relevant (e.g., we omit testing above 90% if failures are already observed at 90%). All statistics are computed over 40 Monte Carlo runs unless mentioned otherwise.

Scale Estimation: Given two point clouds \mathcal{A} and \mathcal{B} , we first create $N(N - 1)/2$ TIMs corresponding to a complete graph and then use Algorithm 2 to solve for the scale. We compute both *Consensus Maximization* [119] and TLS estimates of the scale. Fig. 4(a) shows box plots of the scale error with increasing outlier ratios. The scale error is computed as $|\hat{s} - s^\circ|$, where \hat{s} is the scale estimate and s° is the ground-truth. We observe the TLS solver is robust against 80% outliers, while Consensus Maximization failed twice in that regime.

Maximal Clique Inlier Selection (MCIS): We downsample Bunny to $N = 1000$ and fix the scale to $s = 1$ when applying the random transformation. We first prune the outlier TIMs/TRIMs (edges) that are not consistent with the scale $s = 1$, while

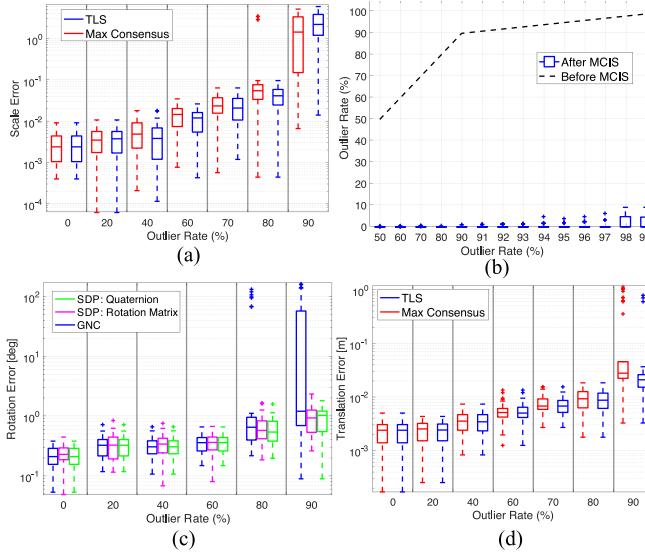


Fig. 4. Results for scale, rotation, translation estimation, and impact of MCIS for increasing outlier rates. (a) Scale estimation. (b) MCIS. (c) Rotation estimation. (d) Translation estimation.

keeping all the points (vertices), to obtain the graph $\mathcal{G}(\mathcal{V}, \mathcal{E}')$. Then we compute the *maximum clique* in $\mathcal{G}(\mathcal{V}, \mathcal{E}')$ and remove all edges and vertices outside the maximum clique, obtaining a pruned graph $\mathcal{G}'(\mathcal{V}', \mathcal{E}'')$ (*cf.* line 1 in Algorithm 1). Fig. 4(b) shows the outlier ratio in \mathcal{G} (label: “Before MCIS”) and \mathcal{G}' (label: “After MCIS”). MCIS effectively reduces outlier rate to below 10%, facilitating rotation and translation estimation, which, in isolation, can already tolerate more than 90% outliers (<90% when using GNC).

Rotation Estimation: We simulate TIMs by applying a random rotation \mathbf{R} to the Bunny, and fixing $s = 1$ and $t = 0$ (we set the number of TIMs to $K = 40$). We compare three approaches to solve the TLS rotation estimation problem (12): 1) the quaternion-based SDP relaxation in (23) and [40] (SDP: Quaternion), 2) the rotation-matrix-based SDP relaxation we proposed in [39] (SDP: Rotation Matrix), and 3) the GNC heuristic in [35]. For each approach, we evaluate the rotation error as $|\arccos((\text{tr}(\hat{\mathbf{R}}^\top \mathbf{R}^\circ) - 1)/2)|$, which is the geodesic distance between the rotation estimate $\hat{\mathbf{R}}$ (produced by each approach) and the ground-truth \mathbf{R}° [120].

Fig. 4(c) reports the rotation error for increasing outlier rates. The GNC heuristic performs well below 80% outliers and then starts failing. The two relaxations ensure similar performance, while the quaternion-based relaxation proposed in this article is slightly more accurate at high outlier rates. Results in Appendix R in [27] show that the quaternion-based relaxation is always tighter than the relaxation in [39], which often translates into better estimates. The reader can find more experiments using our quaternion-based relaxation in [40], where we also demonstrate its robustness against over 95% outliers and discuss applications to panorama stitching.

While the quaternion-based relaxation dominates in terms of accuracy and robustness, it requires solving a large SDP. Therefore, in TEASER++, we opted to use the fast GNC heuristic instead: this choice is motivated by the observation that MCIS is already able to remove most of the outliers [see Fig. 4(b)]

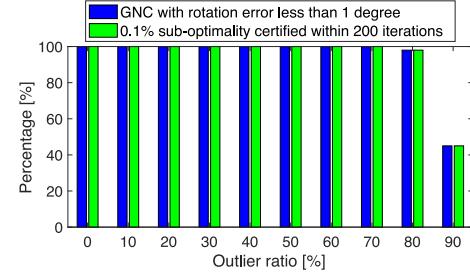


Fig. 5. GNC and optimality certification for TLS robust rotation search.

hence within TEASER++ the rotation estimation only requires solving a problem with less than 10% outliers.

Translation Solver: We apply a random translation \mathbf{t} to the Bunny, and fix $s = 1$ and $\mathbf{R} = \mathbf{I}_3$. Fig. 4(d) shows that component-wise translation estimation using both Consensus Maximization and TLS are robust against 80% outliers. The translation error is defined as $\|\hat{\mathbf{t}} - \mathbf{t}^\circ\|$, the two-norm of the difference between the estimate $\hat{\mathbf{t}}$ and the ground-truth \mathbf{t}° . As mentioned above, most outliers are typically removed before translation estimation, hence enabling TEASER and TEASER++ to be robust to extreme outlier rates (more in Section XI-B).

Optimality Certification for Rotation Estimation: We now test the effectiveness of Algorithm 3 in certifying optimality of a rotation estimate. We consider the same setup, we used for testing rotation estimation but with $K = 100$ TIMs. We use the GNC scheme in [35] to solve problem (12) and then use Algorithm 3 to certify the solution of GNC. Fig. 5 shows the performance of GNC and the certification algorithm under increasing outlier rates, where 100 Monte Carlo runs are performed at each outlier rate. The blue bars show the percentage of runs for which GNC produced a solution with less than 1° rotation error with respect to the ground truth; the green bars show the percentage of runs for which Algorithm 3 produced a relative suboptimality bound lower than 0.1% in less than 200 iterations of DRS. Fig. 5 demonstrates that: 1) the GNC scheme typically produces accurate solutions to the TLS rotation search problem with <80% outliers [a result that confirms the errors we observed in Fig. 4(c)]; 2) the optimality certification algorithm 3 can certify *all* the correct GNC solutions and reject *all* incorrect GNC solutions within 200 iterations. On average, the certification algorithm takes 24 iterations to obtain <0.1% suboptimality bound, where each DRS iteration takes 50 ms in C++.

B. Benchmarking on Standard Datasets

Testing Setup: We benchmark TEASER and TEASER++ against two state-of-the-art robust registration techniques: FGR [56] and *Guaranteed Outlier REMoval (GORE)* [19]. In addition, we test two RANSAC variants (with 99% confidence): a fast version where we terminate RANSAC after a maximum of 1000 iterations (RANSAC1K) and a slow version where we terminate RANSAC after 60 s (RANSAC1min). We use four datasets, Bunny, Armadillo, Dragon, and Buddha, from the Stanford 3-D Scanning Repository [98] and downsample them to $N = 100$ points. The tests below follow the same protocol of Section XI-A. Here, we focus on the results on the Bunny dataset and we postpone the (qualitatively similar) results obtained on the other three datasets to Appendix S in [27].

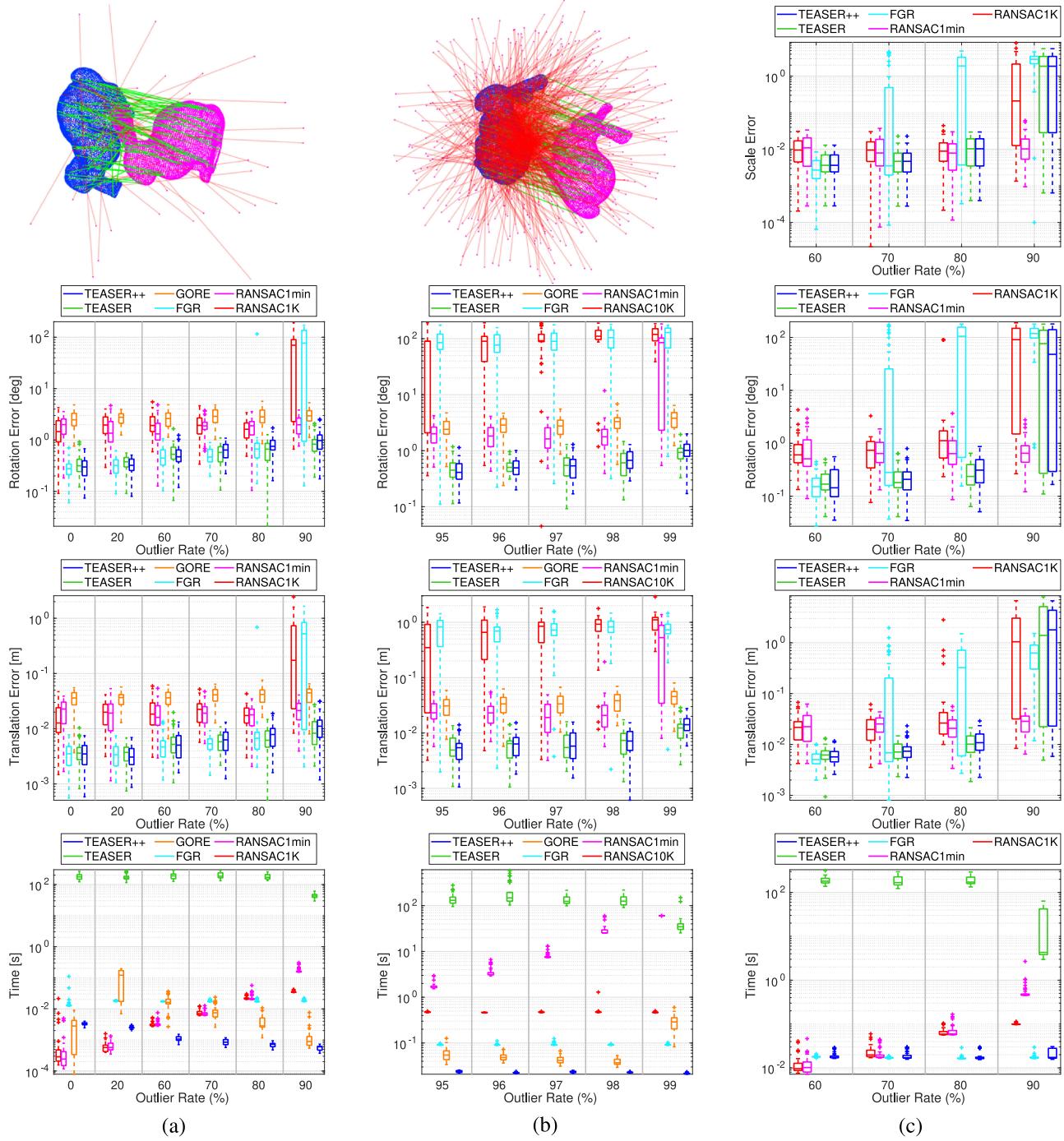


Fig. 6. Benchmark results. (a) Boxplots of rotation errors, translation errors, and timing for the six compared methods on the Bunny dataset with known scale (the top figure shows a registration example with 50% outlier correspondences). (b) Same as before, but for outlier rates between 95% and 99% (the top figure shows an example with 95% outlier correspondences). (c) Boxplots of scale, rotation, translation errors and timing for the five registration methods that support scale estimation on the Bunny dataset with unknown scale. (a) Known scale. (b) Known scale (Extreme Outliers). (c) Unknown scale.

Appendix S also showcases the proposed algorithms on registration problems with high noise ($\sigma = 0.1$), and with up to 10 000 correspondences.

Known Scale: We first evaluate the compared techniques with known scale $s = 1$. Fig. 6(a) shows the rotation error, translation error, and timing for increasing outlier rates on the Bunny dataset. From the rotation and translation errors, we note that TEASER, TEASER++, GORE, and RANSAC1min are robust

against up to 90% outliers, although TEASER and TEASER++ tend to produce more accurate estimates than GORE, and RANSAC1min typically requires over 10^5 iterations for convergence at 90% outlier rate. FGR can only resist 70% outliers and RANSAC1K starts breaking at 90% outliers. TEASER++'s performance is on par with TEASER for all outlier rates, which is expected from the observations in Section X. The timing subplot at the bottom of Fig. 6(a) shows that TEASER is impractical for

real-time robotics applications. On the other hand, TEASER++ is one of the fastest techniques across the spectrum, and is able to solve problems with large number of outliers in less than 10 ms on a standard laptop.

Extreme Outlier Rates: We further benchmark the performance of TEASER and TEASER++ under extreme outlier rates from 95% to 99% with known scale and $N = 1000$ correspondences on the Bunny. We replace RANSAC1K with RANSAC10K, since RANSAC1K already performs poorly at 90% outliers. Fig. 6(b) shows the boxplots of the rotation errors, translation errors, and timing. TEASER, TEASER++, and GORE are robust against up to 99% outliers, while RANSAC1min with 60 s timeout can resist 98% outliers with about 10^6 iterations. RANSAC10K and FGR perform poorly under extreme outlier rates. While GORE, TEASER, and TEASER++ are both robust against 99% outliers, TEASER, and TEASER++ produce lower estimation errors, with TEASER++ being one order of magnitude faster than GORE (bottom sub-figure). We remark that TEASER++’s robustness against 99% outliers is due in large part to the drastic reduction of outlier rates by MCIS.

Unknown Scale: GORE is unable to solve for the scale, hence we only benchmark TEASER and TEASER++ against FGR,⁹ RANSAC1K, and RANSAC1min. Fig. 6(c) plots scale, rotation, translation error and timing for increasing outliers on the Bunny dataset. All the compared techniques perform well when the outlier ratio is below 60%. FGR has the lowest breakdown point and fails at 70%. RANSAC1K, TEASER, and TEASER++ only fail at 90% outlier ratio when the scale is unknown. Although RANSAC1min with 60 s timeout outperforms other methods at 90% outliers, it typically requires more than 10^5 iterations to converge, which is not practical for real-time applications. TEASER++ consistently runs in less than 30 ms.

C. Simultaneous Pose and Correspondences (SPC)

Here, we provide a proof-of-concept of how to use TEASER++ in the case, where we do not have correspondences.

Testing Setup: We obtain the source point cloud \mathcal{A} by down-sampling the Bunny dataset to 100 points. Then, we create the point cloud \mathcal{B} by applying a random rotation and translation to \mathcal{A} . Finally, we remove a percentage of the points in \mathcal{B} to simulate *partial overlap* between \mathcal{A} and \mathcal{B} . For instance, when the overlap is 80%, we discard 20% (randomly chosen) points from \mathcal{B} . We compare TEASER++ against 1) ICP initialized with the identity transformation; and 2) Go-ICP [20] with 30%, 60%, and 90% trimming percentages to be robust to partial overlap. For TEASER++, we generate *all possible correspondences*: in other words, for each point in \mathcal{A} we add all points in \mathcal{B} as a potential correspondence (for a total of $|\mathcal{A}| \cdot |\mathcal{B}|$ correspondences). We then feed the correspondences to TEASER++ that computes a transformation without the need for an initial guess. Clearly, most of the correspondences fed to TEASER are outliers, but we rely on TEASER++ to find the small set of inliers.

Fig. 7(a) and (b) shows the rotation and translation errors for different levels of overlap between \mathcal{A} and \mathcal{B} . ICP fails to compute the correct transformation in practically all instances, since the initial guess is not in the basin of convergence of the optimal solution. Go-ICP is a global method which does

⁹Although the original algorithm in [56] did not solve for the scale, we extend it by using Horn’s method to compute the scale at each iteration.

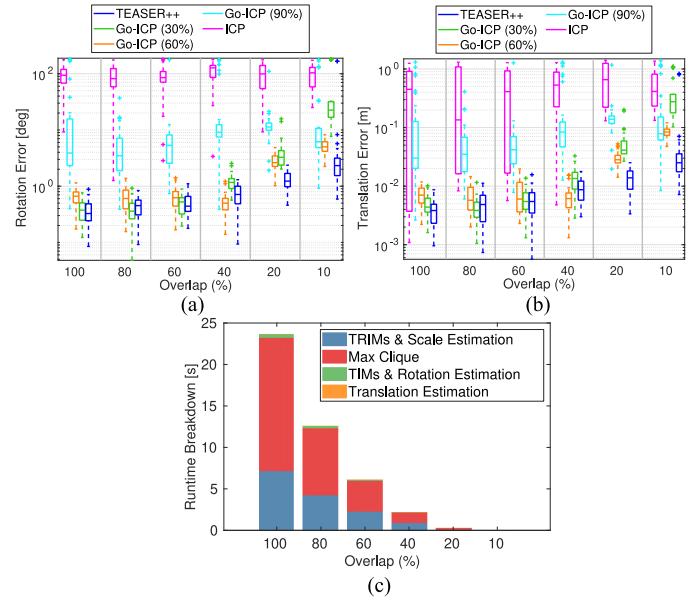


Fig. 7. (a)–(b) Rotation and translation errors for TEASER++, ICP, and Go-ICP in a correspondence-free problem. (c) Timing breakdown for TEASER++. (a) Rotation error. (b) Translation error. (c) Timing breakdown.

not require an initial guess and performs much better than ICP. However, Go-ICP requires the user to set a trimming percentage to deal with partial overlap, and Fig. 7(a) and (b) shows the performance of Go-ICP is highly sensitive to the trimming percentage (Go-ICP (60%) performs the best but is still less accurate than TEASER++). Moreover, Go-ICP takes 16 s on average due to its usage of BnB. TEASER++ computes a correct solution across the spectrum without the need for an initial guess. TEASER++ only starts failing when the overlap drops below 10%. The price we pay for this enhanced robustness is an increase in runtime. We feed $|\mathcal{A}| \cdot |\mathcal{B}| \approx 10^4$ correspondences to TEASER++, which increases the runtime compared to the correspondence-based setup in which the number of correspondences scales linearly (rather than quadratically) with the point cloud size. Fig. 7(c) reports the timing breakdown for the different modules in TEASER++. From the figure, we observe that (i) for small overlap, TEASER++ is not far from real-time, (ii) the timing is dominated by the maximum clique computation and scale estimation, where the latter includes the computation of the TRIMs, and (iii) TEASER++ is orders of magnitude faster than Go-ICP when the overlap is low (e.g., below 40%).

D. Application 1: Object Pose Estimation and Localization

Testing Setup: We use the large-scale point cloud datasets from [37] to test TEASER in *object pose estimation and localization* applications. We first use the ground-truth object labels to extract the *cereal box/cap* out of the scene and treat it as the object, then apply a random transformation to the scene, to get an object-scene pair. To register the object-scene pair, we first use FPFH feature descriptors [1] to establish putative correspondences. Given correspondences from FPFH, TEASER is used to find the relative pose between the object and scene. We downsample the object and scene using the same ratio (about 0.1) to make the object have 2000 points.



Fig. 8. Successful object pose estimation by TEASER on a real RGB-D dataset. Blue lines are the original FPFH [1] correspondences with outliers, green lines are the inlier correspondences computed by TEASER, and the final registered object is highlighted in red.

TABLE I
REGISTRATION RESULTS ON EIGHT SCENES OF THE RGB-D DATASET [37]

	Mean	SD
Rotation error [rad]	0.066	0.043
Translation error [m]	0.069	0.053
# of FPFH correspondences	525	161
FPFH inlier ratio [%]	6.53	4.59

Results: Fig. 8 shows the noisy FPFH correspondences, the inlier correspondences obtained by TEASER, and successful localization and pose estimation of the *cereal box*. Another example is given in Fig. 1(g) and (h). Qualitative results for eight scenes are given in Appendix T in [27]. The inlier correspondence ratios for *cereal box* are all below 10% and typically below 5%. TEASER is able to compute a highly accurate estimate of the pose using a handful of inliers. Table I shows the mean and standard deviation (SD) of the rotation and translation errors, the number of FPFH correspondences, and the inlier ratio estimated by TEASER on the eight scenes.

E. Application 2: Scan Matching

Testing Setup: TEASER++ can also be used in robotics applications that need robust scan matching, such as 3-D reconstruction and loop closure detection in SLAM [121]. We evaluate TEASER++’s performance in such scenarios using the 3DMatch dataset [38], which consists of RGB-D scans from 62 real-world indoor scenes. The dataset is divided into 54 scenes for training, and 8 scenes for testing. The authors provide 5000 randomly sampled keypoints for each scan. On average, there are 205 pairs of scans per scene (maximum: 519 in the Kitchen scene, minimum: 54 in the Hotel 3 scene).

We use 3DSmoothNet [2], a state-of-the-art neural network, to compute local descriptors for each 3-D keypoint, and generate correspondences using nearest-neighbor search. We then feed the correspondences to TEASER++ and RANSAC (implemented in Open3D [122]) and compare their performance in terms of percentage of successfully matched scans and runtime. Due to the large number of pairs, we need to test, we run the experiments on a server with a Xeon Platinum 8259CL CPU at 2.50 GHz, and allocate 12 threads for each algorithm under test. Two scans are successful matched when the transformation computed by a technique has 1) rotation error smaller than 10° , and 2) translation error less than 30 cm. We report RANSAC’s results with maximum number of iterations equal to 1000 (RANSAC1K) and 10 000 (RANSAC10K). We also compare the percentage of successful registrations out of the cases, where TEASER++ certified the rotation estimation as optimal, a setup we refer to as TEASER++ (CERT). The latter executes Algorithm 3 to evaluate the poses from TEASER++ and

TABLE II
PERCENTAGE OF CORRECT REGISTRATION RESULTS AND AVERAGE RUNTIME
USING TEASER++, TEASER++ CERTIFIED, AND RANSAC ON THE
3DMATCH DATASET

	Scenes							Avg. Runtime [s]
	Kitchen (%)	Home 1 (%)	Home 2 (%)	Hotel 1 (%)	Hotel 2 (%)	Hotel 3 (%)	Study (%)	
RANSAC-1K	91.3	89.1	74.5	94.2	84.6	90.7	86.3	81.8
RANSAC-10K	97.2	92.3	79.3	96.5	86.5	94.4	90.4	85.7
TEASER++	98.6	92.9	86.5	97.8	89.4	94.4	91.1	0.059
TEASER++ (CERT)	99.4	94.1	88.7	98.2	91.9	94.4	94.3	238.136

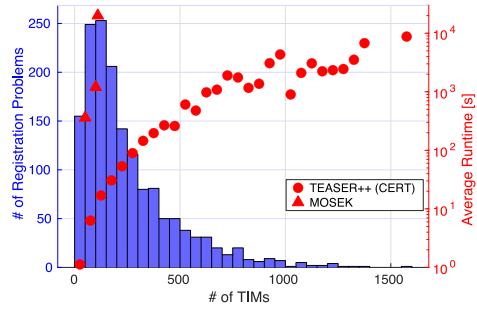


Fig. 9. # of registration problems (histogram) and average runtime (scatter plot, compared with MOSEK) for TEASER++ (CERT) w.r.t. # of TIMs passed to certification in *3DMatch*.

uses a desired suboptimality gap $\bar{\eta} = 3\%$ to certify optimality. We use $\beta_i = \beta = 5$ cm for TEASER++ in all tests.

Results: Table II shows the percentage of successfully matched scans and the average timing for the four compared techniques. TEASER dominates both RANSAC variants with exception of the Lab scene. RANSAC1K has a success rate up to 12% worse than TEASER++. RANSAC10K is an optimized C++ implementation and, while running slower than TEASER++, it cannot reach the same accuracy for most scenes. These results further highlight that TEASER++ can be safely used as a faster and more robust replacement for RANSAC in SLAM pipelines. This conclusion is further reinforced by the last row in Table II, where we show the success rate for the poses certified as optimal by TEASER++. The success rate strictly dominates both RANSAC variants and TEASER++, since TEASER++ (CERT) is able to identify and reject unreliable registration results. This is a useful feature when scan matching is used for loop closure detection in SLAM, since bad registration results lead to incorrect loop closures and can compromise the quality of the resulting map (see [35], [55]). While running Algorithm 3 in TEASER++ (CERT) requires more than 200 s in average, the majority of instances are solved by TEASER++ (CERT) within 100 s and problems with fewer than 50 TIMs can be certified in 1.12 s. Fig. 9 compares the runtime of our certification approach against MOSEK (which directly solves the SDP relaxation) for 50, 100, and 110 TIMs. We can see that TEASER++ (CERT) is orders of magnitude faster, and can certify large-scale problems beyond the reach of MOSEK, which runs out of memory for over 150 TIMs.

Why is TEASER++ not able to solve 100% of the tests in Table II? This should not come as a surprise from the statements

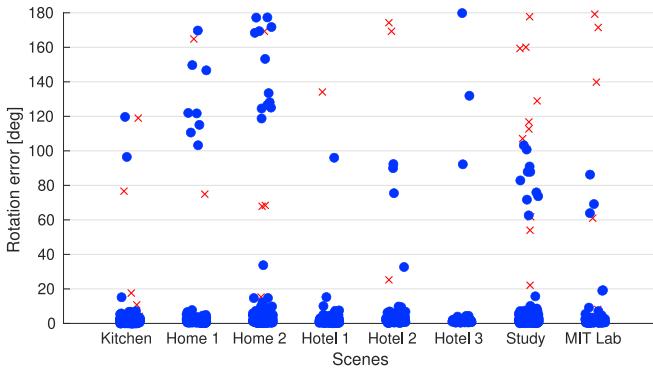


Fig. 10. Rotation errors for each scene (data points correspond to pairs of scans in the scene), with certified TEASER++ solutions (blue dots) versus noncertified (red crosses).

in Theorems 15–17. The estimation contracts discussed in Section IX require a minimum number of inliers: TEASER++ can mine a small number of inliers among a large number of outliers, but it cannot solve problems, where no inliers are given or where the inliers are not enough to identify a unique registration! For the experiments in this section, it is not uncommon to have no or fewer than three inliers in a scene due to the quality of the keypoint descriptors. Moreover, it is not uncommon to have symmetries in the scene, which make the registration non-unique. To intuitively highlight this issue, Fig. 10 shows the rotation errors of TEASER++, with different markers for certified (blue dots) and noncertified (red crosses) solutions. The figure shows that 1) TEASER++ (CERT) is able to reject a large number of incorrect estimates (red crosses with large errors), and 2) some of the incorrect but certified solutions exhibit errors around 90° and 180° which correspond to symmetries of the scene. A visual example of the phenomenon is shown in Fig. 11. These results highlight the need for better keypoint detectors, since even state-of-the-art deep-learning-based methods struggle to produce acceptable outlier rates in real problems.

XII. CONCLUSION

In this article, we propose the first fast and certifiable algorithm for correspondence-based registration with extreme outlier rates. We leverage insights from estimation theory (e.g., unknown-but-bounded noise), geometry (e.g., invariant measurements), graph theory (e.g., maximum clique for inlier selection), and optimization (e.g., tight SDP relaxations). These insights lead to the design of two certifiable registration algorithms. TEASER is accurate and robust but requires solving a large SDP. TEASER++ has similar performance in practice but circumvents the need to solve an SPD and can run in milliseconds. TEASER++ is also certifiable by leveraging DRS to compute a dual optimality certificate. For both algorithms, we provide theoretical bounds on the estimation errors, which are the first of their kind for robust registration problems. Moreover, we test their performance on standard benchmarks, object detection datasets, and the 3DMatch scan matching dataset and show that 1) both algorithms dominate the state of the art (e.g., RANSAC, BnB, heuristics) and are robust to more than 99% outliers, 2) TEASER++ can run in milliseconds and it is currently the fastest robust registration algorithm at high outlier rates, 3) TEASER++ is so robust it can also solve problems without correspondences

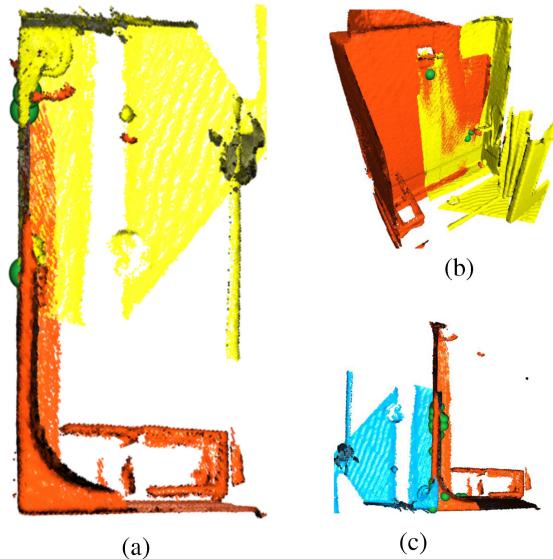


Fig. 11. Example from the Hotel 3 scene of the 3DMatch dataset [38]. The scene pictures a hotel shower where symmetries in the keypoint distribution cause TEASER++ to fail. (a)-(b) Ground-truth alignment (top view and side view). The ground-truth alignment admits 3 inlier correspondences (in green). (c) TEASER++’s estimate of the alignment. TEASER++ is able to find an estimate with 57 inliers. The symmetries of the scene allow for multiple potential registrations, and the ground truth inliers are not part of the maximum clique of the TIM graph. In this case, the max clique selected by TEASER++ contains an incorrect solution that differs from the ground truth by 180°. (a) Ground truth (top-down view). (b) Ground truth (side view). (c) TEASER++’s estimate (top-down view).

(e.g., hypothesizing all-to-all correspondences), where it outperforms ICP and Go-ICP. We release a fast open-source C++ implementation of TEASER++.

While not central to the technical contribution, this article also establishes the foundations of *certifiable perception*, as discussed in Appendix A in [27]. This research area offers many research opportunities and the resulting progress has the potential to boost trustworthiness and reliability in safety-critical application of robotics and computer vision. Future work includes developing certifiable algorithms for other spatial perception problems (ideally leading to certifiable approaches for all the applications discussed in [35] and more). Another avenue for future research is the use of the certifiably robust algorithms presented in this article for self-supervision of deep learning methods for keypoint detection and matching.

ACKNOWLEDGMENT

The authors would like to thank the Associate Editor and the anonymous reviewers for their constructive feedback, and Álvaro Parra and Russ Tedrake for providing comments on an early draft of this article.

REFERENCES

- [1] R. Rusu, N. Blodow, and M. Beetz, “Fast point feature histograms (FPFH) for 3d registration,” in *Proc. IEEE Int. Conf. Robot. Autom.*, 2009, pp. 3212–3217.
- [2] Z. Gojcic, C. Zhou, J. D. Wegner, and A. Wieser, “The perfect match: 3d point cloud matching with smoothed densities,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 5545–5554.

- [3] P. Henry, M. Krainin, E. Herbst, X. Ren, and D. Fox, "RGB-D mapping: Using kinect-style depth cameras for dense 3d modeling of indoor environments," *Int. J. Robot. Res.*, vol. 31, no. 5, pp. 647–663, 2012.
- [4] G. Blais and M. D. Levine, "Registering multiview range data to create 3d computer objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 8, pp. 820–824, Aug. 1995.
- [5] S. Choi, Q. Y. Zhou, and V. Koltun, "Robust reconstruction of indoor scenes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 5556–5565.
- [6] J. Zhang and S. Singh, "Visual-lidar odometry and mapping: Low-drift, robust, and fast," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2015, pp. 2174–2181.
- [7] B. Drost, M. Ulrich, N. Navab, and S. Ilic, "Model globally, match locally: Efficient and robust 3D object recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 998–1005.
- [8] J. M. Wong et al., "SegICP: Integrated deep semantic segmentation and pose estimation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2017, pp. 5784–5789.
- [9] A. Zeng et al., "Multi-view self-supervised deep learning for 6d pose estimation in the amazon picking challenge," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2017, pp. 1386–1383.
- [10] P. Marion, P. R. Florence, L. Manuelli, and R. Tedrake, "Label fusion: A pipeline for generating ground truth labels for real RGBD data of cluttered scenes," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2018, pp. 1–8.
- [11] J. Bazin, Y. Seo, R. Hartley, and M. Pollefeys, "Globally optimal inlier set maximization with unknown rotation and focal length," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 803–817.
- [12] M. A. Audette, F. P. Ferrie, and T. M. Peters, "An algorithmic overview of surface registration techniques for medical imaging," *Med. Image Anal.*, vol. 4, no. 3, pp. 201–217, 2000.
- [13] G. K. L. Tam et al., "Registration of 3d point clouds and meshes: A survey from rigid to nonrigid," *IEEE Trans. Vis. Comput. Graph.*, vol. 19, no. 7, pp. 1199–1217, Jul. 2013.
- [14] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Amer.*, vol. 4, no. 4, pp. 629–642, Apr. 1987.
- [15] K. Arun, T. Huang, and S. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp. 698–700, Sep. 1987.
- [16] F. Tombari, S. Salti, and L. D. Stefano, "Performance evaluation of 3d keypoint detectors," *Intl. J. Comput. Vis.*, vol. 102, no. 1–3, pp. 198–220, 2013.
- [17] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.
- [18] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography," *Commun. ACM*, vol. 24, pp. 381–395, 1981.
- [19] Á. Parra Bustos and T. J. Chin, "Guaranteed outlier removal for point cloud registration with correspondences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 12, pp. 2868–2882, Dec. 2018.
- [20] J. Yang, H. Li, D. Campbell, and Y. Jia, "Go-ICP: A globally optimal solution to 3D ICP point-set registration," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 11, pp. 2241–2254, Nov. 2016.
- [21] A. Bandeira, "A note on probably certifiably correct algorithms," *Comptes Rendus Mathématique*, vol. 354, no. 3, pp. 329–333, 2016.
- [22] L. Carlone and F. Dellaert, "Duality-based verification techniques for 2D SLAM," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2015, pp. 4589–4596. [Online]. Available: [https://www.dropbox.com/s/4wtxypl7hdfnna/2015c-ICRA-duality2D.pdf?dl=0\(pdf\); https://www.bitbucket.org/lucacarlone/pgod2-duality-opencode\(code\)](https://www.dropbox.com/s/4wtxypl7hdfnna/2015c-ICRA-duality2D.pdf?dl=0(pdf); https://www.bitbucket.org/lucacarlone/pgod2-duality-opencode(code))
- [23] D. Rosen, L. Carlone, A. Bandeira, and J. Leonard, "SE-Sync: A certifiably correct algorithm for synchronization over the Special Euclidean group," *Int. J. Robot. Res.*, vol. 38, nos. 2/3, pp. 95–125, 2019.
- [24] Q. M. Rahman, N. Sünderhauf, and F. Dayoub, "Did you miss the sign? A false negative alarm system for traffic sign detectors," *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2019.
- [25] S. Seshia and D. Sadigh, "Towards verified artificial intelligence," 2016, *arXiv: 1606.08514*.
- [26] A. Desai, T. Dreossi, and S. Seshia, "Combining model checking and runtime verification for safe robotics," in *Proc. Int. Conf. Runtime Veri.*, 2017, pp. 172–189.
- [27] H. Yang, J. Shi, and L. Carlone, "TEASER: Fast and certifiable point cloud registration," 2020, *arXiv: 2001.07715*.
- [28] A. Makadia, A. Patterson, and K. Daniilidis, "Fully automatic registration of 3d point clouds," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1, 2006, pp. 1297–1304.
- [29] Y. Liu, C. Wang, Z. Song, and M. Wang, "Efficient global point cloud registration by matching rotation invariant features through translation search," in *Proc. Eur. Conf. Comput. Vis.*, Sep. 2018, pp. 448–463.
- [30] M. Milanese, *Estimation and Prediction in the Presence of Unknown but Bounded Uncertainty: A Survey*. Boston, MA, USA: Springer, 1989, pp. 3–24.
- [31] D. Bertsekas, "Control of uncertain systems with a set-membership description of the uncertainty," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, USA, 1971.
- [32] R. Hartley and F. Kahl, "Global optimization through rotation space search," *Int. J. Comput. Vis.*, vol. 82, no. 1, pp. 64–79, 2009.
- [33] G. Wahba, "A least squares estimate of satellite attitude," *SIAM Rev.*, vol. 7, no. 3, pp. 409–409, 1965.
- [34] I. Diakonikolas, G. Kamath, D. Kane, J. Li, A. Moitra, and A. Stewart, "Robust estimators in high dimensions without the computational intractability," in *Proc. IEEE 57th Annu. Symp. Found. Comput. Sci.*, 2016, pp. 655–664.
- [35] 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 Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1127–1134, Apr. 2020.
- [36] H. Yang and L. Carlone, "One ring to rule them all: Certifiably robust geometric perception with outliers," in *Proc. Adv. Neural Inform. Process. Syst.*, 2020.
- [37] K. Lai, L. Bo, X. Ren, and D. Fox, "A large-scale hierarchical multi-view RGB-D object dataset," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2011, pp. 1817–1824.
- [38] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, and T. Funkhouser, "3dmatch: Learning the matching of local 3d geometry in range scans," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, paper 4.
- [39] H. Yang and L. Carlone, "A polynomial-time solution for robust registration with extreme outlier rates," in *Proc. Robot.: Sci. Syst.*, 2019. [Online]. Available: [http://rss2019.informatik.uni-freiburg.de/papers/0013_FI.pdf\(pdf\); http://rss2019.informatik.uni-freiburg.de/videos/0013_VI_fi.mp4\(video\); http://news.mit.edu/2019/spotting-objects-cars-robots-0620\(media\); https://www.sciencedaily.com/releases/2019/06/190620121444.htm\(media\); http://www.ansa.it/canale_scienza_tecnica/notizie/tecnologie/2019/06/21/i-robot-imparano-a-vedere-nella-nebbia-_9e59485c-ff17-4d62-8224-1d42f44111b9.html\(media\)](http://rss2019.informatik.uni-freiburg.de/papers/0013_FI.pdf(pdf); http://rss2019.informatik.uni-freiburg.de/videos/0013_VI_fi.mp4(video); http://news.mit.edu/2019/spotting-objects-cars-robots-0620(media); https://www.sciencedaily.com/releases/2019/06/190620121444.htm(media); http://www.ansa.it/canale_scienza_tecnica/notizie/tecnologie/2019/06/21/i-robot-imparano-a-vedere-nella-nebbia-_9e59485c-ff17-4d62-8224-1d42f44111b9.html(media))
- [40] H. Yang and L. Carlone, "A quaternion-based certifiably optimal solution to the Wahba problem with outliers," in *Proc. Int. Conf. Comput. Vis. (ICCV)*, 2019.
- [41] C. Choy, J. Park, and V. Koltun, "Fully convolutional geometric features," in *Proc. Int. Conf. Comput. Vis.*, 2019, pp. 8958–8966.
- [42] C. Olsson, F. Kahl, and M. Oskarsson, "Branch-and-bound methods for Euclidean registration problems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 5, pp. 783–794, May 2009.
- [43] J. Briales and J. Gonzalez-Jimenez, "Convex global 3D registration with Lagrangian duality," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 5612–5621.
- [44] F. L. Markley and J. L. Crassidis, *Fundamentals of Spacecraft Attitude Determination and Control*. New York, NY, USA: Springer, 2014, vol. 33.
- [45] F. L. Markley, "Attitude determination using vector observations and the singular value decomposition," *J. Astronaut. Sci.*, vol. 36, no. 3, pp. 245–258, 1988.
- [46] J. Gower and G. Dijksterhuis, "Procrustes problems," in *Procrustes Problems*, Oxford Statistical Science Series, vol. 30, Oxford, U.K.: Oxford Univ. Press, 2005.
- [47] P. Schonemann, "A generalized solution of the orthogonal procrustes problem," *Psychometrika*, vol. 31, pp. 1–10, 1966.
- [48] N. Ohta and K. Kanatani, "Optimal estimation of three-dimensional rotation and reliability evaluation," *IEICE Trans. Inform. Syst.*, vol. 81, no. 11, pp. 1247–1252, 1998.
- [49] Y. Cheng and J. L. Crassidis, "A total least-squares estimate for attitude determination," in *Proc. AIAA Scitech Forum*, 2019, paper 1176.
- [50] S. Ahmed, E. C. Kerrigan, and I. M. Jaïmoukha, "A semidefinite relaxation-based algorithm for robust attitude estimation," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 3942–3952, Aug. 2012.
- [51] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

- [52] P. Meer, D. Mintz, A. Rosenfeld, and D. Y. Kim, "Robust regression methods for computer vision: A review," *Int. J. Comput. Vis.*, vol. 6, no. 1, pp. 59–70, Apr. 1991.
- [53] K. M. Tavish and T. D. Barfoot, "At all costs: A comparison of robust cost functions for camera correspondence outliers," in *Proc. IEEE 12th Conf. Comput. Robot Vis.*, 2015, pp. 62–69.
- [54] M. J. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision," *Int. J. Comput. Vision*, vol. 19, no. 1, pp. 57–91, 1996.
- [55] P. Lajoie, S. Hu, G. Beltrame, and L. Carlone, "Modeling perceptual aliasing in SLAM via discrete-continuous graphical models," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1232–1239, Apr. 2019.
- [56] Q. Zhou, J. Park, and V. Koltun, "Fast global registration," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 766–782.
- [57] O. Enqvist, K. Josephson, and F. Kahl, "Optimal correspondences from pairwise constraints," in *Proc. Int. Conf. Comput. Vis.*, 2009, pp. 1295–1302.
- [58] A. Parra Bustos, T.-J. Chin, F. Neumann, T. Friedrich, and M. Katzmann, "A practical maximum clique algorithm for matching with pairwise constraints," 2019, *arXiv:1902.01534*.
- [59] V. Golyanik, S. Aziz Ali, and D. Stricker, "Gravitational approach for point set registration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 5802–5810.
- [60] P. Jauer, I. Kuhleemann, R. Bruder, A. Schweikard, and F. Ernst, "Efficient registration of high-resolution feature enhanced point clouds," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 5, pp. 1102–1115, May 2019.
- [61] H. Yang, "A dynamical perspective on point cloud registration," 2020, *arXiv:2005.03190*.
- [62] T. J. Chin and D. Suter, "The maximum consensus problem: Recent algorithmic advances," *Synthesis Lectures Comput. Vis.*, vol. 7, no. 2, pp. 1–194, 2017.
- [63] F. Wen, R. Ying, Z. Gong, and P. Liu, "Efficient algorithms for maximum consensus robust fitting," *IEEE Trans. Robot.*, vol. 36, no. 1, pp. 92–106, Feb. 2020.
- [64] Z. Cai, T.-J. Chin, and V. Koltun, "Consensus maximization tree search revisited," in *Proc. Int. Conf. Comput. Vis.*, 2019, pp. 1637–1645.
- [65] H. M. Le, T.-J. Chin, A. Eriksson, T.-T. Do, and D. Suter, "Deterministic approximate methods for maximum consensus robust fitting," *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI)*, 2019.
- [66] V. Tzoumas, P. Antonante, and L. Carlone, "Outlier-robust spatial perception: Hardness, general-purpose algorithms, and guarantees," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2019, pp. 5383–5390.
- [67] T.-J. Chin, Z. Cai, and F. Neumann, "Robust fitting in computer vision: Easy or hard?," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 701–716.
- [68] D. Campbell, L. Petersson, L. Kneip, and H. Li, "Globally-optimal inlier set maximisation for simultaneous camera pose and feature correspondence," in *Proc. Int. Conf. Comput. Vis.*, 2017, pp. 1–10.
- [69] J. C. Bazin, Y. Seo, and M. Pollefeys, "Globally optimal consensus set maximization through rotation search," in *Proc. Asian Conf. Comput. Vis.*, 2012, pp. 539–551.
- [70] C. Olsson, O. Enqvist, and F. Kahl, "A polynomial-time bound for matching and registration with outliers," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.
- [71] O. Enqvist, E. Ask, F. Kahl, and K. Åström, "Robust fitting for multiple view geometry," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 738–751.
- [72] E. Ask, O. Enqvist, and F. Kahl, "Optimal geometric fitting under the truncated L_2 -norm," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2013, pp. 1722–1729.
- [73] S. Granger and X. Pennec, "Multi-scale EM-ICP: A fast and robust approach for surface registration," in *Proc. Eur. Conf. Comput. Vis.*, 2002, pp. 418–432.
- [74] L. Maier-Hein *et al.*, "Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 8, pp. 1520–1532, Aug. 2012.
- [75] D. Chetverikov, D. Stepanov, and P. Krsek, "Robust Euclidean alignment of 3D point sets: The trimmed iterative closest point algorithm," *Image Vis. Comput.*, vol. 23, no. 3, pp. 299–309, 2005.
- [76] S. Kaneko, T. Kondo, and A. Miyamoto, "Robust matching of 3D contours using iterative closest point algorithm improved by M-estimation," *Pattern Recognit.*, vol. 36, no. 9, pp. 2041–2047, 2003.
- [77] A. Myronenko and X. Song, "Point set registration: Coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.
- [78] B. Jian and B. C. Vemuri, "Robust point set registration using gaussian mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1633–1645, Aug. 2011.
- [79] W. Clark, M. Ghaffari, and A. Bloch, "Nonparametric continuous sensor registration," 2020, *arXiv:2001.04286*.
- [80] H. M. Le, T.-T. Do, T. Hoang, and N.-M. Cheung, "SDRSAC: Semidefinite-based randomized approach for robust point cloud registration without correspondences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 124–133.
- [81] T. M. Breuel, "Implementation techniques for geometric branch-and-bound matching methods," *Comput. Vis. Image Understanding*, vol. 90, no. 3, pp. 258–294, 2003.
- [82] Á. Parra Bustos, T. J. Chin, and D. Suter, "Fast rotation search with stereographic projections for 3d registration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2014, pp. 3930–3937.
- [83] H. Li and R. Hartley, "The 3D-3D registration problem revisited," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2007, pp. 1–8.
- [84] G. Izatt, H. Dai, and R. Tedrake, "Globally optimal object pose estimation in point clouds with mixed-integer programming," *Robot. Res.*, 2020, pp. 695–710.
- [85] H. Maron, N. Dym, I. Kezurer, S. Kovalsky, and Y. Lipman, "Point registration via efficient convex relaxation," *ACM Trans. Graph.*, vol. 35, no. 4, pp. 1–12, 2016.
- [86] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3D classification and segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 652–660.
- [87] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph CNN for learning on point clouds," *ACM Trans. Graph.*, vol. 38, no. 5, 2019, Art. no. 146.
- [88] Y. Aoki, H. Goforth, R. A. Srivatsan, and S. Lucey, "Pointnetlk: Robust & efficient point cloud registration using PointNet," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 7163–7172.
- [89] Y. Wang and J. M. Solomon, "Deep closest point: Learning representations for point cloud registration," in *Proc. Int. Conf. Comput. Vis.*, 2019, pp. 3523–3532.
- [90] Y. Wang and J. M. Solomon, "PRNet: Self-supervised learning for partial-to-partial registration," in *Proc. Adv. Neural Inform. Process. Syst.*, 2019, pp. 8812–8824.
- [91] A. Avetisyan, M. Dahnert, A. Dai, M. Savva, A. X. Chang, and M. Nießner, "Scan2CAD: Learning CAD model alignment in RGB-D scans," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 2614–2623.
- [92] A. Avetisyan, A. Dai, and M. Nießner, "End-to-End CAD model retrieval and 9DoF alignment in 3D scans," in *Proc. Int. Conf. Comput. Vis.*, 2019, pp. 2551–2560.
- [93] C. Choy, W. Dong, and V. Koltun, "Deep global registration," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 2511–2520.
- [94] J. Fredriksson, V. Larsson, C. Olsson, and F. Kahl, "Optimal relative pose with unknown correspondences," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 1728–1736.
- [95] T.-Y. Liu and H. Jiang, "Minimizing sum of truncated convex functions and its applications," *J. Comput. Graph. Statist.*, vol. 28, no. 1, pp. 1–10, 2019.
- [96] S. Barratt, G. Angeris, and S. Boyd, "Minimizing a sum of clipped convex functions," *Optim. Lett.*, pp. 1–17, 2020.
- [97] F. Chung, *Spectral Graph Theory*. American Mathematical Soc., CBMS Regional Conference Series in Mathematics, No. 92, 1996.
- [98] B. Curless and M. Levoy, "A volumetric method for building complex models from range images," in *Proc. SIGGRAPH*, 1996, pp. 303–312.
- [99] X. Li, Y. Liu, Y. Wang, C. Wang, M. Wang, and Z. Song, "Fast and globally optimal rigid registration of 3d point sets by transformation decomposition," 2019, *arXiv:1812.11307*.
- [100] S. Agarwal, V. Shree, and S. Chakravorty, "RFM-SLAM: Exploiting relative feature measurements to separate orientation and position estimation in slam," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2017, pp. 6307–6314.
- [101] L. Carlone, R. Aragues, J. Castellanos, and B. Bona, "A fast and accurate approximation for planar pose graph optimization," *Int. J. Robot. Res.*, vol. 33, no. 7, pp. 965–987, 2014. [Online]. Available: <https://www.dropbox.com/s/z1ol1o17zye5qxz/2014j-IJRR-LAGO.pdf?dl=0> (pdf) <https://www.dropbox.com/s/bojshzsdd4q8m6at/2011-RSS-presentation.pdf?dl=0> (ppt) <https://bitbucket.org/gtborg/>

- gtsam(code) <https://lucacarlone.mit.edu/media/video> (datasets: [https://lucacarlone.mit.edu/datasets/\(web\)](https://lucacarlone.mit.edu/datasets/(web)))
- [102] S. Perera and N. Barnes, "Maximal cliques based rigid body motion segmentation with a RGB-D camera," in *Proc. Asian Conf. Comput. Vision*, 2012, pp. 120–133.
- [103] C. Bron and J. Kerbosch, "Algorithm 457: Finding all cliques of an undirected graph," *Commun. ACM*, vol. 16, no. 9, pp. 575–577, 1973.
- [104] B. Pattabiraman, M. M. A. Patwary, A. H. Gebremedhin, W. K. Liao, and A. Choudhary, "Fast algorithms for the maximum clique problem on massive graphs with applications to overlapping community detection," *Internet Math.*, vol. 11, no. 4–5, pp. 421–448, 2015.
- [105] Q. Wu and J. Hao, "A review on algorithms for maximum clique problems," *Eur. J. Oper. Res.*, vol. 242, no. 3, pp. 693–709, 2015.
- [106] D. Eppstein, M. Löffler, and D. Strash, "Listing all maximal cliques in sparse graphs in near-optimal time," in *Proc. Int. Symp. Algorithms Comput.*, 2010, pp. 403–414.
- [107] D. Scaramuzza, "1-point-ransac structure from motion for vehicle-mounted cameras by exploiting non-holonomic constraints," *Int. J. Comput. Vis.*, vol. 95, no. 1, pp. 74–85, 2011.
- [108] M. ApS, *The MOSEK Optimization Toolbox for MATLAB Manual. Version 8.1.*, 2017. [Online]. Available: <http://docs.mosek.com/8.1/toolbox/index.html>
- [109] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [110] D. Henrion and J. Malick, "Projection methods in conic optimization," in *Handbook on Semidefinite, Conic and Polynomial Optimization*. New York, NY, USA: Springer, 2012, pp. 565–600.
- [111] H. H. Bauschke and J. M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Rev.*, vol. 38, no. 3, pp. 367–426, 1996.
- [112] N. J. Higham, "Computing a nearest symmetric positive semidefinite matrix," *Linear Algebra Appl.*, vol. 103, pp. 103–118, 1988.
- [113] P. L. Combettes and J.-C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*. New York, NY, USA: Springer, 2011, pp. 185–212.
- [114] S. Jegelka, F. Bach, and S. Sra, "Reflection methods for user-friendly submodular optimization," in *Proc. Adv. Neural Inform. Process. Syst.*, 2013, pp. 1313–1321.
- [115] OpenMP Architecture Review Board, "OpenMP application program interface version 3.0," May 2008. [Online]. Available: <http://www.openmp.org/mp-documents/spec30.pdf>
- [116] R. A. Rossi, D. F. Gleich, and A. H. Gebremedhin, "Parallel maximum clique algorithms with applications to network analysis," *SIAM J. Scientific Comput.*, vol. 37, no. 5, pp. C589–C616, 2015.
- [117] M. Quigley *et al.*, "Ros: An open-source robot operating system," in *Proc. ICRA Workshop Open Source Softw.*, 2009, paper 5.
- [118] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," 2014. [Online]. Available: <http://cvxr.com/cvx>
- [119] P. Speciale, D. P. Paudel, M. R. Oswald, T. Kroeger, L. V. Gool, and M. Pollefeys, "Consensus maximization with linear matrix inequality constraints," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 5048–5056.
- [120] R. Hartley, J. Trumpf, Y. Dai, and H. Li, "Rotation averaging," *IJCV*, vol. 103, no. 3, pp. 267–305, 2013.
- [121] C. Cadena *et al.*, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, Dec. 2016.
- [122] Q.-Y. Zhou, J. Park, and V. Koltun, "Open3D: A modern library for 3D data processing," 2018, *arXiv:1801.09847*.
- [123] S. Jegelka and J. Bilmes, "Submodularity beyond submodular energies: Coupling edges in graph cuts," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 1897–1904.
- [124] J.-B. Lasserre, *Moments, Positive Polynomials and Their Applications*. Singapore: World Scientific, 2010, vol. 1.



Heng Yang received the B.S. degree (Hons.) in mechanical engineering from Tsinghua University, Beijing, China, in 2015, and the M.S. degree in mechanical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2017. He is currently working toward the Ph.D. degree in robotics, computer vision, optimization and learning with the Department of Mechanical Engineering and the Laboratory for Information & Decision Systems (LIDS), Massachusetts Institute of Technology.

He is currently working with Prof. Luca Carlone at the SPARK Lab. His work includes developing robust and certifiable algorithms for geometric understanding. His research interests include convex optimization, semidefinite and sums-of-squares relaxation, robust estimation and machine learning, applied to robotic perception and computer vision.

Mr. Yang is a recipient of the Best Paper Award in Robot Vision at the 2020 IEEE International Conference on Robotics and Automation (ICRA).



Jingnan Shi received the B.S. degree in engineering from Harvey Mudd College, Claremont, CA, USA, in 2019. He is currently working toward the master's degree in robotics autonomy with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA.

He is currently a member of the SPARK Lab, led by Prof. Luca Carlone. His research interests include robust perception and estimation, with applications specific to various robotic systems.



Luca Carlone received the B.S. degree in mechatronics from the Polytechnic University of Turin, Turin, Italy, in 2006, the M.S. degree in mechatronics from the Polytechnic University of Turin, in 2008, the M.S. degree in automation engineering from the Polytechnic University of Milan, Milan, Italy, in 2008, and the Ph.D. degree in robotics from the Polytechnic University of Turin, in 2012.

He is currently the Leonardo CD Assistant Professor with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, and a Principal Investigator with the Laboratory for Information & Decision Systems (LIDS). He joined LIDS as a Postdoctoral Associate in 2015 and later as a Research Scientist in 2016, after spending two years as a Postdoctoral Fellow with the Georgia Institute of Technology from 2013 to 2015. His work includes seminal results on certifiably correct algorithms for localization and mapping, as well as approaches for visual-inertial navigation and distributed mapping. His research interests include nonlinear estimation, numerical and distributed optimization, and probabilistic inference, applied to sensing, perception, and decision-making in single and multirobot systems.

Prof. Carlone is a recipient of the 2020 RSS Early Career Award, the 2017 Transactions on Robotics King-Sun Fu Memorial Best Paper Award, the Best Paper Award in Robot Vision at ICRA 2020, the Best Paper award at WAFR 2016, and the Best Student Paper Award at the 2018 Symposium on VLSI Circuits.