

Semidefinite Relaxations for Robust Multiview Triangulation

Linus Härenstam-Nielsen^{1,2}, Niclas Zeller⁴, Daniel Cremers^{1,2,3}

¹Technical University of Munich, ²Munich Center for Machine Learning, ³University of Oxford,
⁴Karlsruhe University of Applied Sciences

linus.nielsen@tum.de, niclas.zeller@h-ka.de, cremers@tum.de

Abstract

We propose an approach based on convex relaxations for certifiably optimal robust multiview triangulation. To this end, we extend existing relaxation approaches to non-robust multiview triangulation by incorporating a least squares cost function. We propose two formulations, one based on epipolar constraints and one based on fractional reprojection constraints. The first is lower dimensional and remains tight under moderate noise and outlier levels, while the second is higher dimensional and therefore slower but remains tight even under extreme noise and outlier levels. We demonstrate through extensive experiments that the proposed approaches allow us to compute provably optimal reconstructions even under significant noise and a large percentage of outliers.

1. Introduction

Multiview triangulation is the problem of estimating the location of a point in 3D given two or more 2D observations in images taken from cameras with known poses and intrinsics. The 2D observations are typically estimated by some form of feature matching pipeline, so they are always corrupted by noise and outliers. As a result the 3D point cannot be exactly recovered, and instead the solution has to be phrased as a nonconvex optimization problem.

While solutions are typically computed using faster but sub-optimal local optimization methods, there have also been efforts to compute globally optimal triangulations using semidefinite relaxations [1, 4, 13]. These relaxations can work well even in high-noise scenarios, but their practical use remains limited as they are not robust and even a single outlier can deteriorate the result significantly. In this work, inspired by recent advances in semidefinite relaxations for outlier-robust perception [28], we will show that [1, 4] can be extended to also handle significant amounts of outliers.

Implementation: github.com/linusnie/robust-triangulation-relaxations

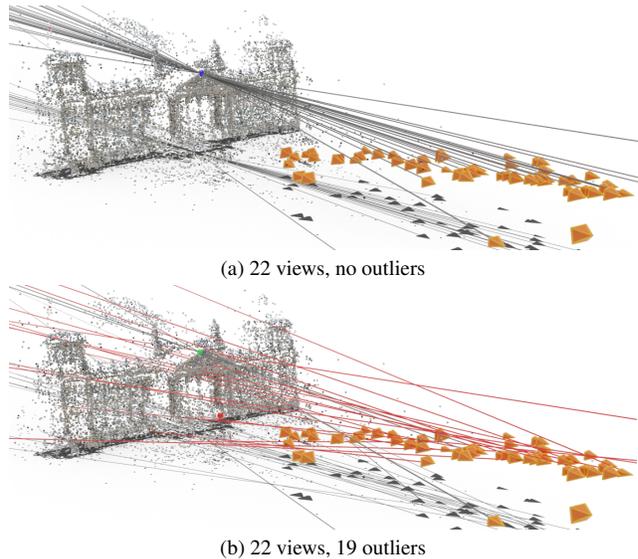


Figure 1. Example of a triangulated point from the Reichstag dataset. **Blue point**: ground truth from [12]. **Red point**: non-robust global optimum found by the relaxation from [1] (see Eq. (T)). **Green point**: robust global optimum found by our proposed relaxation in Eq. (RT).

Semidefinite relaxations have the advantage of being globally solvable in polynomial time, meaning that they can be used to enable practical *certifiably optimal* algorithms. After solving the relaxed problem we either have that 1) the relaxation is tight and we provably recover the global optimum of the original problem, or 2) the relaxation is provably not tight and we can report failure to find the global optimum. The key metric for the usefulness of a certifiably optimal algorithm is then the percentage of problem cases where the underlying relaxation is tight.

Despite their often slower runtime, certifiably optimal methods offer several advantages: Firstly, in safety-critical systems it may be required or desirable to complement the computed solution with some guarantee that the solver is not stuck in a local optima. Secondly, in many offline applications runtime is actually not as critical and then one may want to trade off better accuracy for extra runtime. Thirdly, globally optimal solutions of real-world problems can serve

as ground truth for assessing the performance of local optimization methods.

In this work, we demonstrate a certifiably optimal approach to robust triangulation by developing two convex relaxations for the truncated least squares cost function. Enabling the combination of robustness with the capacity to compute certifiably optimal solutions. Our main contributions can be summarized as follows:

- We extend the convex triangulation methods from [1] and [4] with a truncated least squares cost function and propose two corresponding convex relaxations.
- We validate empirically that both relaxations remain tight even under large amounts of noise and high outlier ratios.
- We show that the relaxations are tight in the noise-free and outlier-free case by explicitly constructing the dual solution.

To the best of our knowledge, this is the first example of a successful semidefinite relaxation of a robust estimation problem with reprojection errors.

2. Related work

Triangulation is a core subroutine for structure from motion and therefore has been studied extensively. For two views, there are many globally optimal solution variants, including computing the roots of a degree 6 polynomial [9] for the reprojection error or a 3×3 singular value decomposition for the angular error (up to a second order approximation) [16]. Outlier-free multiview triangulation is arguably most commonly solved based on the linear-eigen method from [9]. Robust triangulation is typically tackled using RANSAC [14, 17, 21] where a 2-view solver is repeatedly applied to randomly sampled pairs of views until an inlier set can be established.

Semidefinite relaxations have been used to obtain certifiably optimal algorithms for many problems in geometric computer vision. Examples include semidefinite relaxations for partitioning, grouping and restoration [15], for minimizing reprojection errors [13], for multiview triangulation [1, 4], for essential matrix estimation [30], for hand-eye calibration [7, 24, 25], for robust point cloud registration [26, 28, 29], and for 3D shape from 2D landmarks [27].

Notably [28], which is one of the main inspirations for this work, demonstrates that semidefinite relaxations can also be used for outlier-robust estimation. In particular, they provide relaxations for various outliers models in the context of robust rotation averaging, mesh registration, absolute pose registration and category-level object pose+shape estimation.

Solving semidefinite relaxations is typically slow and memory intensive, stemming from the fact that the number

of variables is the square of the number of variables in the original problem. To tackle this issue, there has been recent interest in developing solvers that can scale to larger problems. Including [6] which uses a reformulation in terms of eigenvalue optimization based on [10] which can take advantage of GPUs, and [28] which uses efficient non-global solvers for speeding up the convergence of the global solver. Notably, in both cases the main memory saving comes from applying a dual-only solver.

In a limited number of cases, semidefinite relaxations can be shown to always find a globally optimal solution to the original problem. This includes the dual quaternion formulation of hand-eye calibration [7] and 2-view triangulation using epipolar constraints [1], in both cases with some assumption of non-degenerate measurements. Another example is the rotation alignment problem which has a closed form solution in terms of an eigenvalue decomposition (quaternion formulation) or singular value decomposition (rotation matrix formulation).

However, outlier-robust estimation is inapproximable in general [2], meaning there will always be some subset of possible measurements for which any semidefinite relaxation of practical size is non-tight. In terms of theoretical guarantees, [5] introduces the concept of *local stability* which, under certain conditions on the problem structure for noise-free measurements, can guarantee that a relaxation remains tight for bounded measurement noise. In some cases it is also possible to find conditions on measurements which guarantee that the relaxation is tight or non-tight, as demonstrated in [20] for robust rotation alignment of point clouds, but typically algorithm developers will have to rely on experiments in order to determine to what extent a given relaxation remains tight in a particular problem scenario.

3. Notation and preliminaries

For $t, s \in \mathbb{R}^3$ we write $[t]_{\times}$ for the 3×3 skew-symmetric matrix such that $t \times s = [t]_{\times} s$. $(a; b)$ denotes the vertical concatenation of vectors a and b and for a collection of vectors a_1, \dots, a_n the subscript-free version denotes the corresponding stacked vector $a = (a_1; \dots; a_n)$. We use a bar to denote the *homogeneous* version of a vector, that is $\bar{a} := (a; 1)$. When dimensionality is understood we define e_i to be the i th unit vector and $E_i = e_i e_i^T$. For a vector of monomials $m = (m_1; \dots; m_d)$ we define $e_{m_i}^m$ as the unit vector whose only non-zero entry corresponds to the index of m_i in m , meaning $e_{m_i}^m = e_i \in \mathbb{R}^d$. For a vector $x \in \mathbb{R}^k$ we define:

$$M_x := \begin{pmatrix} I & -x \\ -x^T & \|x\|^2 \end{pmatrix} \in \mathbb{S}_{k+1} \quad (1)$$

such that for $y \in \mathbb{R}^k$ we have $\bar{y}^T M_x \bar{y} = \|x - y\|^2$. The operator \otimes denotes the Kronecker product, and \oplus denotes

the tensor sum. For example, for 2×2 matrices A and B :

$$A \oplus B = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}, A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B \\ a_{21}B & a_{22}B \end{pmatrix}. \quad (2)$$

3.1. Semidefinite relaxations

As a general strategy, we aim to solve the triangulation problem by relaxing a *Quadratically Constrained Quadratic Program* (QCQP) which has the following form:

$$\begin{aligned} \min_{z \in \mathbb{R}^d} \quad & z^T M z \\ \text{s.t.} \quad & z^T E z = 1 \\ & z^T A_i z = 0, \quad i = 1, \dots, k \\ & z^T B_j z \leq 0, \quad j = 1, \dots, l. \end{aligned} \quad (3)$$

This is a very general formulation with applications in computer vision but it is NP-hard to solve in most cases, so an imperfect method is typically necessary. One such strategy is to lift the problem from \mathbb{R}^d to the set of $d \times d$ positive semidefinite matrices, \mathbb{S}_d , by introducing a new variable $Z = z z^T$ and using the fact that $z^T M z = \text{tr}(M z z^T) = \text{tr}(M Z)$ to arrive at:

$$\begin{aligned} \min_{Z \in \mathbb{S}_d} \quad & \text{tr}(M Z) \\ \text{s.t.} \quad & \text{tr}(E Z) = 1 \\ & \text{tr}(A_i Z) = 0, \quad i = 1, \dots, k \\ & \text{tr}(B_j Z) \leq 0 \quad j = 1, \dots, l \\ & Z \succeq 0. \end{aligned} \quad (4)$$

Eq. (4) is a relaxation of Eq. (3) since if z satisfies the constraints of Eq. (3) we always have that $Z = z z^T$ satisfies the constraints of Eq. (4) with the same objective value. However, the converse is not always true. In particular, if \hat{Z} is optimal for Eq. (4) we can obtain a corresponding solution \hat{z} for Eq. (3) with the same objective value if and only if \hat{Z} is rank one. In this case we have $\hat{Z} = \hat{z} \hat{z}^T$ and we then say that the relaxation is *tight*.

The main advantage of working with the relaxation Eq. (4) as opposed to the original problem Eq. (3) is that the relaxation is a convex optimization problem, in particular it is a semidefinite program, for which a variety of polynomial-time solvers are available, including [3, 18]. If the relaxation is not tight we can at best expect an optimal \hat{Z} to generate an approximation of the optimal \hat{z} . Therefore, a key metric to consider when applying a relaxation is the percentage of encountered problem cases in which it remains tight.

4. Relaxations for multiview triangulation

Given n views of a point X from cameras located at $P_i = (R_i, t_i) \in \text{SE}(3)$ in camera-to-world convention with in-

trinsic matrices $K_i \in \mathbb{R}^{3 \times 3}$, and with, possibly noisy, observations denoted as $\tilde{x}_i \in \mathbb{R}^2$, the n -view triangulation problem with reprojection error is defined as:

$$\min_{X \in \mathbb{R}^3} \sum_{i=1}^n \|\tilde{x}_i - \pi(K_i, P_i, X)\|^2 \quad (5)$$

where $\pi(K_i, P_i, X)$ is the reprojection of the point $X \in \mathbb{R}^3$ to camera i . This is a nonconvex problem but it is not yet in QCQP form since $\pi(K_i, P_i, X)$ is not quadratic in X . In this section we will recap two ways of converting Eq. (5) to a QCQP, from which we can generate the corresponding semidefinite relaxations.

4.1. Triangulation with epipolar constraints

The first approach was introduced in [1]. They showed that Eq. (5) can be formulated as a polynomial optimization problem of degree 2 by reparametrizing X in terms of its n reprojections x_i , which are constrained to satisfy the epipolar constraints:

$$\begin{aligned} \min_{x_i \in \mathbb{R}^2} \quad & \sum_{i=1}^n \|x_i - \tilde{x}_i\|^2 \\ \text{s.t.} \quad & \tilde{x}_i^T F_{ij} \tilde{x}_j = 0 \\ & i, j = 1, \dots, n \quad i \neq j \end{aligned} \quad (6)$$

where $F_{ij} = K_i^{-T} [t_{ij}]_{\times} R_{ij} K_j^{-1}$ is the fundamental matrix corresponding to the relative transformation between poses i and j . Since the estimated reprojections x_i all satisfy the epipolar constraints, the solution of Eq. (5) can in most cases be recovered exactly from Eq. (6) using the linear-eigen method from [9]. An important failure case of this parametrization is that when the camera centers are coplanar it is possible that the solution x_i of Eq. (6) does not correspond to a valid 3D point, see Appendix A for more details and an example.

Using the parametrization $z = (x; 1) = \bar{x}$ the semidefinite relaxation of Eq. (6) is:

$$\begin{aligned} \min_{Z \in \mathbb{S}_{2n+1}} \quad & \text{tr}(M_{\bar{x}} Z) \\ \text{s.t.} \quad & \text{tr}(E_{n+1} Z) = 1 \\ & \text{tr}(\bar{F}_{ij} Z) = 0, \quad i = 1, \dots, k \\ & Z \succeq 0 \end{aligned} \quad (\text{T})$$

where $\bar{F}_{ij} \in \mathbb{S}_{2n+1}$ is defined such that $\bar{x}^T \bar{F}_{ij} \bar{x} = \tilde{x}_i^T F_{ij} \tilde{x}_j$. It was shown in both [1] and [5] that the relaxation Eq. (T) is locally stable with respect to noise as long as the views are not coplanar.

4.2. Triangulation with fraction constraints

As initially proposed in [4], an alternative to Eq. (6) is to explicitly parametrize the 3D point X in homogeneous

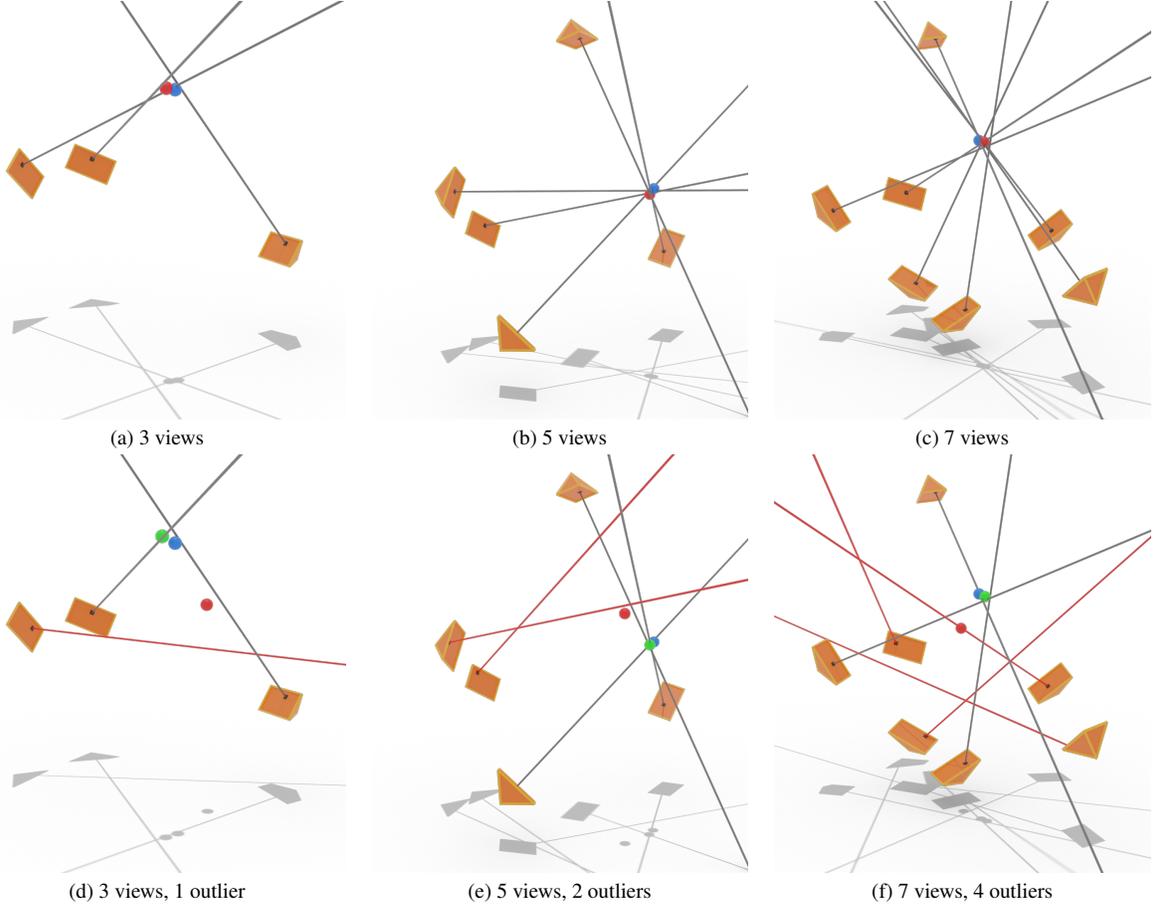


Figure 2. Examples of simulated triangulation problems from Sec. 6.1 with $\sigma = 50\text{px}$ for various number of views and outliers. **Blue point**: ground truth, **Red point**: non-robust global optimum found by the relaxation from [4] (see Eq. (TF)). **Green point**: robust global optimum found by our proposed relaxation in Eq. (RTF). With no outliers the robust and non-robust methods give the same result.

coordinates. This leads to the fractional constraints $x_i^k = a_{ik}^T \bar{X} / b_i^T \bar{X}$, where a_{i1} , a_{i2} and b_i are given by the rows of the i -th camera projection matrix $K_i (R_i^T \quad -R_i^T t_i)$. By multiplying out the right hand side denominators we get the following QCQP:

$$\begin{aligned}
 \min_{\bar{X} \in \mathbb{R}^4, x_i \in \mathbb{R}^2} \quad & \sum_{i=1}^n \|x_i - \tilde{x}_i\|^2 \\
 \text{s.t.} \quad & \bar{X}^T \bar{X} = 1 \\
 & x_i^k b_i^T \bar{X} - a_{ik}^T \bar{X} = 0 \\
 & i = 1, \dots, n \quad k = 1, 2.
 \end{aligned} \tag{7}$$

A naive approach to relaxing Eq. (7) would be to use the parametrization $z = (x; \bar{X})$, but as shown in [4] this leads to a relaxation whose optimal value is always zero. To circumvent this issue, [4] instead proposes parametrizing the problem in terms of all possible products between the elements of x and X , *i.e.* $x \otimes X$. They also show through experiments that, while the resulting relaxation has more parameters and constraints than Eq. (T), it is also tight in

a significantly wider range of cases, leading to a tradeoff between stability and computation time.

We will use a similar relaxation, though we will skip the initial change of variables to get a slightly different but equivalent formulation which can be extended to the robust case more conveniently. We start by setting $z = (x \otimes \bar{X}; \bar{X}) = \bar{x} \otimes \bar{X}$ and then we multiply each reprojection constraint in Eq. (7) with z_j to get $8n + 4$ quadratic constraints:

$$\begin{aligned}
 (x_i^k b_i^T \bar{X} - a_{ik}^T \bar{X}) z_j &= z^T (e_{x_i^k}^{\bar{x}} \otimes b_i - e_1^{\bar{x}} \otimes a_{ik}) e_j^T z \\
 &= 0
 \end{aligned} \tag{8}$$

where we have made use of the unit vector notation from Sec. 3, meaning in particular $e_{x_i^k}^{\bar{x}} = e_{2i+k}$ and $e_1^{\bar{x}} = e_{2n+1}$. We also need to introduce constraints to preserve the fact that z comes from a $(2n+1) \times 4$ Kronecker product. When $Z = zz^T$ is rank one, it turns out that this condition is equivalent to Z being composed of $2n+1$ symmetric 4×4 blocks, see [4] for more details. We will denote this constraint as $Z \in \text{BSym}(2n+1, 4)$. The relaxation of can then

be written as¹:

$$\begin{aligned}
& \min_{Z \in \mathbb{S}_+^{8n+4}} \quad \text{tr}(Z(M_{\bar{x}} \otimes I_4)) \\
\text{s.t.} \quad & \text{tr}(Z(0_{8n \times 8n} \oplus I_4)) = 1 \\
& Z \in \text{BSym}(2n+1, 4) \\
& \text{tr}(Z(e_{x_i^k}^{\bar{x}} \otimes b_i - e_1^{\bar{x}} \otimes a_{ik})e_j^T) = 0 \\
& i = 1, \dots, n, \quad k = 1, 2 \\
& j = 1, \dots, 8n+4.
\end{aligned} \tag{TF}$$

5. The robust case

Now that we have introduced the two main relaxations of Eq. (5) we move to the the main contribution of this paper, which is to introduce the corresponding truncated least squares (TLS) extensions. Similarly to [28] we will use the fact that the TLS cost function can be written as a minimization problem by introducing a binary decision variable for each residual

$$\rho_i(r_i^2) = \min(r_i^2, c_i) = \min_{\theta_i \in \{0,1\}} \theta_i r_i^2 + (1 - \theta_i)c_i \quad (9)$$

where $c_i > 0$ is the square of the inlier threshold. In Appendix B we show that both relaxations are tight in the noise-free and outlier-free case, and we also show part of the criteria required for local stability.

5.1. Robust triangulation with epipolar constraints

Using Eq. (9), the TLS extension of Eq. (6) is:

$$\begin{aligned}
& \min_{x_i \in \mathbb{R}^2, \theta_i \in \mathbb{R}} \quad \sum_{i=1}^n \left(\theta_i \|x_i - \tilde{x}_i\|^2 + (1 - \theta_i)c_i \right) \\
\text{s.t.} \quad & \tilde{x}_i^T F_{ij} \tilde{x}_j = 0, \\
& \theta_i^2 - \theta_i = 0, \quad \sum_{i=1}^n \theta_i^2 \geq 2, \\
& i, j = 1, \dots, n \quad i \neq j.
\end{aligned} \tag{10}$$

where the constraints $\theta_i^2 - \theta_i = 0$ ensures θ_i equals 0 or 1, and the constraint $\sum_{i=1}^n \theta_i^2 \geq 2$ ensures there are at least two inliers in the final solution. This cost function includes terms like $\theta_i \|x_i\|^2$, which means it is a 3rd degree polynomial, so we can't apply the relaxation as described in Sec. 3.1 directly. But we can obtain a 2nd order formulation by noting that $\theta_i^2 = \theta_i$ implies $\theta_i \|x_i - \tilde{x}_i\|^2 = \|\theta_i x_i - \theta_i \tilde{x}_i\|^2$ and making the substitution $y_i = \theta_i x_i$:

$$\begin{aligned}
& \min_{y_i \in \mathbb{R}^2, \theta_i \in \mathbb{R}} \quad \sum_{i=1}^n \left(\|y_i - \theta_i \tilde{x}_i\|^2 + (1 - \theta_i)c_i \right) \\
\text{s.t.} \quad & (y_i; \theta_i)^T F_{ij} (y_j; \theta_j) = 0 \\
& \theta_i^2 - \theta_i = 0, \quad \sum_{i=1}^n \theta_i^2 \geq 2, \\
& \theta_i y_i = y_i \\
& i, j = 1, \dots, n, \quad i \neq j.
\end{aligned} \tag{11}$$

¹The cost functions in Eq. (7) and Eq. (TF) are equivalent, since $(\bar{X} \otimes \bar{x})^T (M_{\bar{x}} \otimes I_4) (\bar{X} \otimes \bar{x}) = (\bar{x}^T M_{\bar{x}} \bar{x}) \bar{X}^T \bar{X} = \bar{x}^T M_{\bar{x}} \bar{x}$.

The last set of constraints $\theta_i y_i = y_i$ are redundant but we find that they are necessary for the relaxation to remain tight in the presence of noise, these are referred to as moment constraints in [28]. We can recover a solution of Eq. (10) from a solution of Eq. (11) by triangulating the estimated inliers (*i.e.* y_i for which $\theta_i = 1$) and setting each x_i (including outliers) to be the reprojection of the resulting point onto view i .

Using the parametrization $z = (y; \theta; 1)$ the semidefinite relaxation of Eq. (11) is:

$$\begin{aligned}
& \min_{Z \in \mathbb{S}_+^{3n+1}} \quad \text{tr}(M_{\bar{x}}^c Z) \\
\text{s.t.} \quad & \text{tr}(\bar{F}_{ij} Z) = 0 \\
& Z_{\theta_i, \theta_i} - Z_{1, \theta_i} = 0 \\
& Z_{\theta_i, y_i} - Z_{1, y_i} = 0 \\
& \text{tr}(E_{3n+1} Z) = 1 \\
& i, j = 1, \dots, n \quad i \neq j
\end{aligned} \tag{RT}$$

where $M_{\bar{x}}^c$ is the robust extension of $M_{\bar{x}}$, defined as:

$$\begin{aligned}
M_{\bar{x}}^c &= \begin{pmatrix} I & -B(\tilde{x}) & 0 \\ -B(\tilde{x})^T & \text{diag}(\|\tilde{x}_i\|^2) & -c \\ 0 & -c^T & \sum_{i=0}^n c_i \end{pmatrix}, \\
B(\tilde{x}) &= \begin{pmatrix} \tilde{x}_1 & 0 & \dots & 0 \\ 0 & \tilde{x}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \tilde{x}_n \end{pmatrix}.
\end{aligned} \tag{12}$$

and Z_{m_i, m_j} is the entry of Z corresponding to the index of the monomials m_i and m_j in z . In Appendix B.1 we analyze the dual of Eq. (RT) to show that the relaxation is tight in the noise-free and outlier-free case.

5.2. Robust triangulation with fraction constraints

In this section we will introduce a higher order relaxation which can handle higher noise and outlier levels. Since the fractional constraints in Eq. (TF) are more stable with respect to noise than the epipolar constraints in Eq. (T), we might also expect that extending Eq. (TF) to handle outliers will result in a relaxation which is more stable than Eq. (RT). In this section we will show how the robust extension can be formulated, and as we will see in Sec. 6 it is indeed significantly more stable with respect to both noise and outliers.

In order to extend Eq. (TF) to handle outliers we will proceed in a similar manner as in the case with epipolar constraints. Starting by writing the cost function in terms of

Problem	Relaxation	Robust	Constraints	Variables
Eq. (6)	Eq. (T)	✗	$\frac{1}{2}n^2 - \frac{1}{2}n + 1$	$2n + 1$
Eq. (11)	Eq. (RT)	✓	$\frac{1}{2}n^2 + 2.5n + 1$	$3n + 1$
Eq. (7)	Eq. (TF)	✗	$28n^2 + 14n + 1$	$8n + 4$
Eq. (13)	Eq. (RTF)	✓	$51n^2 + 65n + 1$	$12n + 4$

Table 1. Summary of relaxations for the triangulation problem and its robust extension.

the 2nd order variables $y_i = \theta_i x_i$:

$$\begin{aligned}
\min_{\bar{X} \in \mathbb{R}^4, x_i \in \mathbb{R}^2} \quad & \sum_{i=1}^n \left(\|y_i - \theta_i \tilde{x}_i\|^2 + (1 - \theta_i) c_i \right) \\
\text{s.t.} \quad & \bar{X}^T \bar{X} = 1 \\
& y_i^k b_i^T \bar{X} - a_{ik}^T \bar{X} = 0 \\
& \theta_i^2 - \theta_i = 0, \quad \sum_{i=1}^n \theta_i^2 \geq 2 \\
& \theta_i y_i = y_i \\
& i = 1, \dots, n \quad k = 1, 2.
\end{aligned} \tag{13}$$

For convenience we will denote the vertical concatenation of y and θ as $(y; \theta) = y_\theta$. For the relaxation we will then use the parametrization $z = (y_\theta \otimes \bar{X}; \bar{X}) = \bar{y}_\theta \otimes \bar{X}$ and generate redundant moment constraints from $\theta_i^2 - \theta_i = 0$ and $\theta_i y_i = y_i$ by multiplying each equation by $\bar{X}_s \bar{X}_t$ for $s, t = 1, \dots, 4$. We also generate redundant inequalities from $\sum_{i=1}^n \theta_i^2 \geq 2$ by multiplying by \bar{X}_s^2 for each $s = 1, \dots, 4$. Resulting in the following relaxation:

$$\begin{aligned}
\min_{Z \in \mathbb{S}_+^{12n+4}} \quad & \text{tr}(Z((M_{\bar{x}}^c \otimes I_4) \oplus 0_{4 \times 4})) \\
\text{s.t.} \quad & \text{tr}(Z(0_{12n \times 12n} \oplus I_4)) = 1 \\
& Z \in \text{BSym}(3n + 1, 4) \\
& \text{tr}(Z(e_{y_i^k}^{\bar{y}_\theta} \otimes b_i - e_{\theta_i}^{\bar{y}_\theta} \otimes a_{ik})e_j^T) = 0 \\
& Z_{\bar{X}_s \theta_i, \bar{X}_t \theta_i} - Z_{\bar{X}_s, \bar{X}_t \theta_i} = 0 \\
& Z_{\bar{X}_s \theta_i, \bar{X}_t y_i} - Z_{\bar{X}_s, \bar{X}_t y_i} = 0 \\
& \sum_{i=1}^n Z_{\theta_i \bar{X}_s, \theta_i \bar{X}_s} \geq 2Z_{\bar{X}_s, \bar{X}_s} \\
& i = 1, \dots, n, \quad k = 1, 2 \quad s, t = 1, \dots, 4 \\
& j = 1, \dots, 12n + 4.
\end{aligned} \tag{RTF}$$

Similarly to the epipolar case we can show that this relaxation is tight in the noise-free and outlier-free case by explicitly constructing the globally optimal Lagrange multipliers, see Appendix B.2 for details.

With this we have 4 relaxations for the triangulation problem corresponding to the non-robust and robust case with the epipolar and the fractional parametrization. We summarize the relaxations and their number of variables and constraints in Tab. 1.

5.3. Rounding in the non-tight case

For non-tight cases the optimal \hat{Z} will have rank of at least 2, which means we can't recover the optimal solution \hat{z} for the original problem Eq. (3). However we can still construct an approximate solution through a rounding procedure. We start by setting \hat{z} to be the eigenvector corresponding to the minimal eigenvalue, normalized such that $\hat{z}^T E \hat{z} = 1$. We then apply a different procedure for each problem depending on the constraints. For Eq. (T) we triangulate the resulting \hat{x}_i (which in this case will generally not satisfy the epipolar constraints) using the linear-eigen method from [9]. For Eq. (RT) we do the same except that we first determine the inlier parameters $\hat{\theta}_i$ by rounding the corresponding entries of \hat{z} to 0 or 1. For Eq. (TF) and Eq. (RTF) we compute the best-fitting tensor product decomposition of \hat{z} using a singular value decomposition as described in [23] and then use the same method as in the epipolar case for determining the inlier parameters.

6. Experiments

We implement all relaxations using CVXPY [8] with the solver MOSEK [3] using the setting `MSK_DPAR_INTPNT_CO_TOL_REL_GAP = 10-14`, all other solver parameters are left on their defaults. We find that working in units of pixels results in poorly conditioned solutions where \hat{z} does not satisfy the constraints to high accuracy even for tight relaxations. To avoid this issue we use the change of variables $x_i \rightarrow \frac{1}{W} x_i$ and adjust the intrinsics accordingly. Since the scaling is the same for each point the optimal solution remains unchanged, but we get much closer to rank one solutions in practice due to the improved numerical stability.

6.1. Simulated experiments

We simulate triangulation problems as initially proposed in [19] by placing n cameras on a sphere of radius 2 and sample a point to be triangulated from the unit cube, see Fig. 2 for some examples. The same setup was also used for experiments in [1, 4]. For the reprojection model we simulate a pinhole camera with parameters from one of the cameras in the Reichstag dataset: width $W = 2108$, height $H = 1162$, focal length $f = 1012.0027$ and principal point $p = (1054, 581)$. We simulate noisy observations by adding Gaussian noise with standard deviation σ to the ground truth image coordinates. When generating an outlier we select a view at random and replace the measurement with a random point in the image.

We run the experiment for each method at various different noise levels and number of outliers. For each noise level we run Eq. (RT) 750 times and Eq. (RTF) 120 times for $n = 3, 5$ and 7 views and in each case add up to $n - 2$ outliers. The percentage of tight relaxations and the estima-

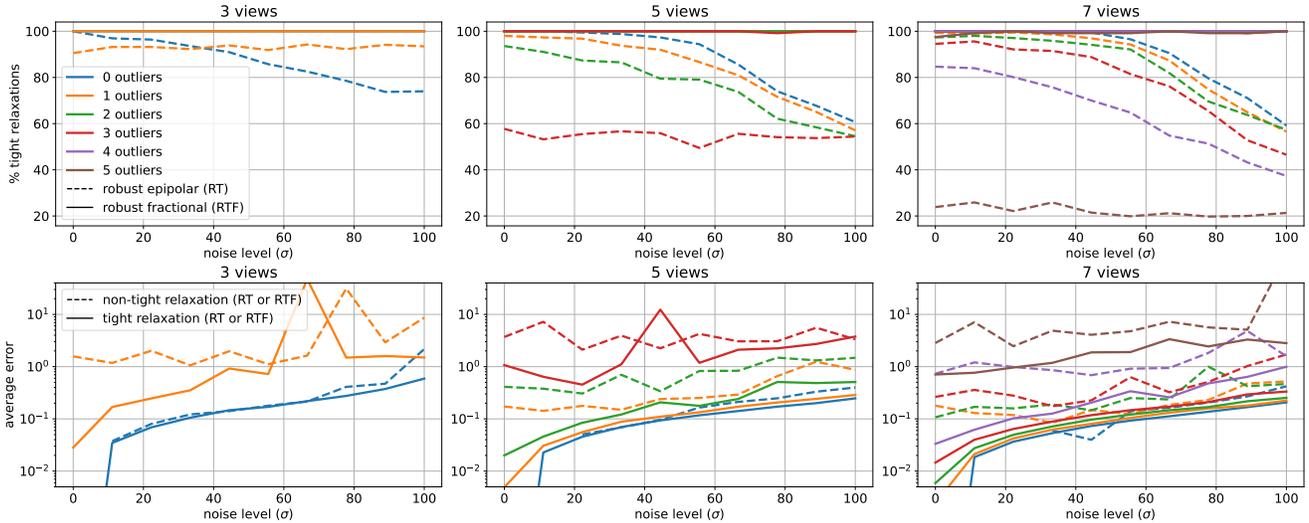


Figure 3. Average number of tight relaxation (top) and estimation error (bottom) for 3, 5 and 7 views for the robust epipolar relaxation Eq. (RT) and the robust fractional relaxation Eq. (RTF) for experiments described in Sec. 6.1. Note that the error plot is divided by tight/non-tight relaxation since both methods give the same result whenever the relaxation is tight.

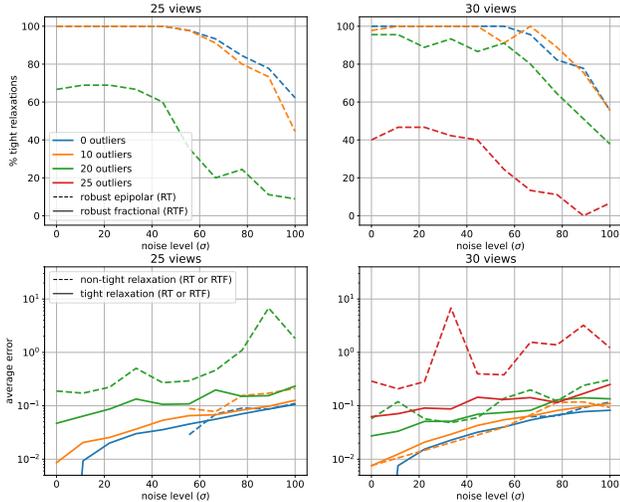


Figure 4. Average number of tight relaxation (top) and estimation error (bottom) for 25 and 30 views using Eq. (RT).

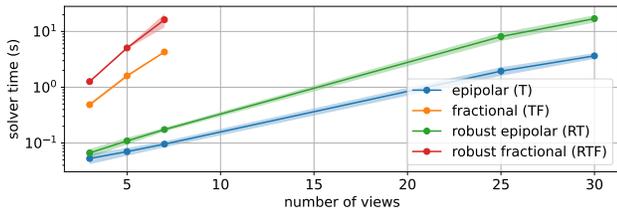


Figure 5. Average computation time for each solver, averaged over all noise levels and number of outliers.

tion error can be seen in Fig. 3. We also run Eq. (RT) 45 times each for $n = 25$ and 30 with 0, 10, 20 and 25 outliers, the results of which can be seen in Fig. 4. We don't run Eq. (RTF) for these cases since we run into memory limitations with MOSEK. We set $c_i = 200^2$ for all experiments.

From Fig. 3 we can see that in general the fractional relaxation in Eq. (13) is significantly more stable than the epipolar relaxation Eq. (RT). In fact, across all experiments

the fractional relaxation is tight in 99.92% of cases. However, we can also note that the epipolar relaxation remains viable for lower noise levels, for instance in the case with $n = 7$ views and 3 outliers the relaxations is tight in more than 90% of cases when the noise is below $\sigma \approx 40$ px, after which the percentage of tight relaxations drop drastically.

As can be seen from the average solver timings in Fig. 5 the fractional relaxations is also over one order of magnitude slower than the epipolar relaxation, meaning that it is preferable to use Eq. (RT) in cases where the quality of observations is known to be high, and fall back to Eq. (RTF) only if the epipolar relaxation is non-tight and spending extra time on reducing the error is desirable.

6.2. Reichstag dataset

We also validate our relaxations on the Reichstag dataset from [12]. The dataset consists of 75 views of roughly 18k 3D points. We use the ground truth correspondences estimated by structure from motion as detailed in [12] and generate each triangulation problem by selecting n views which all observe a common point. We then add up to $n - 2$ outliers by replacing the ground truth observations with a randomly selected keypoints in the same image. See Fig. 1 for an example point with $n = 22$ views and 19 outliers.

For $n = 3, 5$ and 7 views we run Eq. (RT) 375 times and Eq. (RTF) 60 times for each possible number of outliers with $c_i = 10^2$. And similarly we run Eq. (RT) 120 times for $n = 25$ and 30 views. The results are summarized in Fig. 6.

Similarly to the simulated experiments we can note that the percentage of tight relaxations decreases steadily as more outliers are added, with a sharp drop when the number of inliers gets close to 2, with the fractional method outperforming the epipolar method.

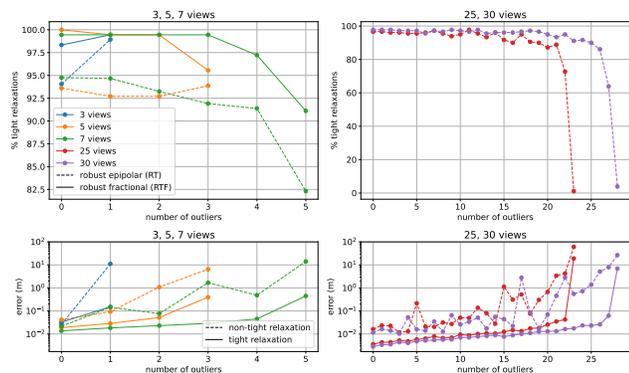


Figure 6. Average number of tight relaxation for Eq. (RT) and Eq. (RTF) (top) and estimation error for tight and non-tight relaxations (bottom) on the Reichstag dataset. See Sec. 6.2 for details.

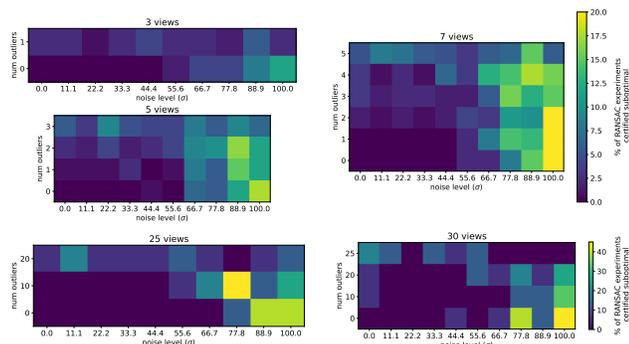


Figure 7. Color indicates the percentage of experiment trials where our robust relaxation (Eq. (RT) for 3, 5 and 7 views and Eq. (RTF) for 25 and 30 views) found a lower cost solution than RANSAC. Details in Sec. 6.3.

6.3. RANSAC comparison

In many scenarios it is unrealistic to spend the computation time necessary for certifiable optimality at runtime, yet certifiably optimal algorithms can still provide valuable insight to algorithm developers in an offline fashion. In this section we will provide an illustrative how this principle can be applied to RANSAC-based triangulation.

In particular, we compare our relaxations with exhaustive MLESAC [22] with TLS objective which we implement in the following way: first off, in order to eliminate the effect of randomness, we evaluate every possible pair of views. For each pair of views we generate a candidate point by computing the optimal triangulation using Eq. (T), which is always tight in the 2-view case (see [1]), and evaluate the robust cost function in Eq. (10). From the candidate point achieving the minimal robust cost we further generate an additional candidate point by locally refining Eq. (5) on the inlier points. The final RANSAC estimate is then whichever of the candidate points obtains the lowest robust cost.

We run our RANSAC implementation on every triangulation problem from the simulated experiments in Sec. 6.1 and compare against the results obtained by the certifiably

views	tight	relaxation better	same solution	RANSAC better
3	✓	69	2331	0
3	✗	0	0	0
5	✓	257	4542	0
5	✗	1	0	0
7	✓	446	6744	0
7	✗	1	6	3

views	tight	relaxation better	same solution	RANSAC better
25	✓	12	319	0
25	✗	32	58	29
30	✓	18	424	0
30	✗	31	84	43

Table 2. Top: Eq. (RT) vs. RANSAC for 3, 5 and 7 views. Bottom: Eq. (RTF) vs. RANSAC for 25 and 30 views.

optimal solvers. In cases where the relaxation is not tight we also refine the rounded solution on the inlier set. We do this for all simulated experiments and compare against our relaxations, see Tab. 2 for a breakdown of the number of cases where RANSAC was certified as optimal/suboptimal. In Fig. 7 we show the percentage of cases where the robust relaxation finds a better solution than RANSAC. For low noise and outlier levels RANSAC more or less always finds the globally optimal solution, while for high noise levels there are many failure cases.

7. Conclusion

We proposed a global optimization framework for robust multiview triangulation. To this end we derive semidefinite relaxations for triangulation losses that incorporate a truncated quadratic cost making them robust to both noise and outliers. On synthetic and real data we confirm that provably optimal triangulations can be computed and relaxations remain empirically tight despite significant amounts of noise and outliers.

7.1. Acknowledgments

This work was supported by the ERC Advanced Grant SIMULACRON, by the Munich Center for Machine Learning and by the EPSRC Programme Grant VisualAI EP/T028572/1.

References

- [1] Chris Aholt, Sameer Agarwal, and Rekha Thomas. A qcqp approach to triangulation. In *European Conference on Computer Vision*, pages 654–667. Springer, 2012. 1, 2, 3, 6, 8, 10
- [2] Pasquale Antonante, Vasileios Tzoumas, Heng Yang, and Luca Carlone. Outlier-robust estimation: Hardness, minimally tuned algorithms, and applications. *IEEE Transactions on Robotics*, 38(1):281–301, 2022. 2
- [3] MOSEK ApS. *The MOSEK optimization toolbox for Python manual. Version 10.0.*, 2022. 3, 6

- [4] Diego Cifuentes. A convex relaxation to compute the nearest structured rank deficient matrix. *SIAM Journal on Matrix Analysis and Applications*, 42(2):708–729, 2021. 1, 2, 3, 4, 6, 11
- [5] Diego Cifuentes, Sameer Agarwal, Pablo A Parrilo, and Rekha R Thomas. On the local stability of semidefinite relaxations. *Mathematical Programming*, 193(2):629–663, 2022. 2, 3, 10, 11
- [6] Sumanth Dathathri, Krishnamurthy Dvijotham, Alexey Kurakin, Aditi Raghunathan, Jonathan Uesato, Rudy R Bunel, Shreya Shankar, Jacob Steinhardt, Ian Goodfellow, Percy S Liang, et al. Enabling certification of verification-agnostic networks via memory-efficient semidefinite programming. *Advances in Neural Information Processing Systems*, 33:5318–5331, 2020. 2
- [7] Amit Dekel, Linus Harenstam-Nielsen, and Sergio Caccamo. Optimal least-squares solution to the hand-eye calibration problem. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 2
- [8] Steven Diamond and Stephen Boyd. CVXPY: A Python-embedded modeling language for convex optimization. *Journal of Machine Learning Research*, 17(83):1–5, 2016. 6
- [9] Richard I. Hartley and Peter Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157, 1997. 2, 3, 6
- [10] Christoph Helmberg and Franz Rendl. A spectral bundle method for semidefinite programming. *SIAM Journal on Optimization*, 10(3):673–696, 2000. 2
- [11] Anders Heyden and Kalle Åström. Algebraic properties of multilinear constraints. *Mathematical Methods in the Applied Sciences*, 20(13):1135–1162, 1997. 10
- [12] Yuhe Jin, Dmytro Mishkin, Anastasiia Mishchuk, Jiri Matas, Pascal Fua, Kwang Moo Yi, and Eduard Trulls. Image Matching across Wide Baselines: From Paper to Practice. *International Journal of Computer Vision*, 2020. 1, 7
- [13] Fredrik Kahl and Didier Henrion. Globally optimal estimates for geometric reconstruction problems. *International Journal of Computer Vision*, 74(1):3–15, 2007. 1, 2
- [14] Lai Kang, Lingda Wu, and Yee-Hong Yang. Robust multi-view L_2 triangulation via optimal inlier selection and 3D structure refinement. *Pattern Recognition*, 47(9):2974–2992, September 2014. 2
- [15] Jens Keuchel, Christoph Schnorr, Christian Schellewald, and Daniel Cremers. Binary partitioning, perceptual grouping, and restoration with semidefinite programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(11):1364–1379, 2003. 2
- [16] Seong Hun Lee and Javier Civera. Closed-form optimal two-view triangulation based on angular errors. pages 2681–2689, 10 2019. 2
- [17] Seong Hun Lee and Javier Civera. Robust uncertainty-aware multiview triangulation. *CoRR*, abs/2008.01258, 2020. 2
- [18] Brendan O’Donoghue, Eric Chu, Neal Parikh, and Stephen Boyd. Conic optimization via operator splitting and homogeneous self-dual embedding. *Journal of Optimization Theory and Applications*, 169(3):1042–1068, June 2016. 3
- [19] Carl Olsson, Fredrik Kahl, and Richard Hartley. Projective least-squares: Global solutions with local optimization. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1216–1223, 2009. 6
- [20] Liangzu Peng, Mahyar Fazlyab, and René Vidal. Semidefinite relaxations of truncated least-squares in robust rotation search: Tight or not. In *European Conference on Computer Vision*, pages 673–691. Springer, 2022. 2
- [21] Johannes Lutz Schönberger and Jan-Michael Frahm. Structure-from-motion revisited. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016. 2
- [22] P.H.S. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156, 2000. 8
- [23] C. F. Van Loan and N. Pitsianis. *Approximation with Kronecker Products*, pages 293–314. Springer Netherlands, Dordrecht, 1993. 6
- [24] Emmett Wise, Matthew Giamou, Soroush Khoubyarian, Abhinav Grover, and Jonathan Kelly. Certifiably optimal monocular hand-eye calibration. In *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 271–278. IEEE, 2020. 2
- [25] Thomas Wodtke, Markus Horn, Michael Buchholz, and Klaus Dietmayer. Globally optimal multi-scale monocular hand-eye calibration using dual quaternions. In *2021 International Conference on 3D Vision (3DV)*, pages 249–257. IEEE, 2021. 2
- [26] Heng Yang and Luca Carlone. A quaternion-based certifiably optimal solution to the wahba problem with outliers. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2019. 2
- [27] Heng Yang and Luca Carlone. In perfect shape: Certifiably optimal 3d shape reconstruction from 2d landmarks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020. 2
- [28] Heng Yang and Luca Carlone. Certifiably optimal outlier-robust geometric perception: Semidefinite relaxations and scalable global optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022. 1, 2, 5
- [29] Heng Yang, Jingnan Shi, and Luca Carlone. Teaser: Fast and certifiable point cloud registration. *IEEE Transactions on Robotics*, 37(2):314–333, 2020. 2
- [30] Ji Zhao. An efficient solution to non-minimal case essential matrix estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020. 2