Computer Vision and Image Understanding
journal homepage: www.elsevier.com

Certifiable algorithms for the two-view planar triangulation problem

Mercedes Garcia-Salguero[a,**], Javier Gonzalez-Jimenez[a]

[a] *Machine Perception and Intelligent Robotics (MAPIR) Group, System Engineering and Automation Department, University of Malaga, Campus de Teatinos 29071 Malaga, Spain*

ABSTRACT

Planar scenes predominate in man-made environments, *e.g.* interior or facades of buildings and in ground images from aerial vehicles. Points lying on those surfaces can be reconstructed from their observations in two images. However, generic reconstruction algorithms output 3D points not lying on the plane, thus obtaining inaccurate reconstructions. The problem also turns to be non-convex with many local minima, hence hindering the performance of iterative method. Therefore, being able to obtain *and* certify the optimal solution to this problem is of special relevant for these applications. In this paper we first propose a fast and certifiable algorithm that both estimates and certifies the optimal solution to the triangulation problem. From this formulation, we also present an optimality certificate that tells us whether a given solution (obtained by any solver) is the global optimum. Last, from this certificate we derive a sufficient (but not necessary) optimality condition that allows us to certify optimality in less than one microsecond. We test the proposed algorithms on extensive experiments on both synthetic and real data. Code is made available at `https://github.com/mergarsal`.

1. Introduction

The triangulation problem consists of estimating the 3D structure of the scene from a set of (potentially) noisy observations and known projective matrices (pose and calibration parameters). This task is at the core of more complex applications, among others Simultaneous Localization and Mapping (SLAM) or Structure-from-Motion (SfM) see *e.g.* Kang et al. (2020); Artieda et al. (2009); Karrer and Chli (2020); Gomez-Ojeda et al. (2019); Mur-Artal et al. (2015); Triggs et al. (1999); Arndt et al.. Previous works have approached this problem for a generic scene, see *e.g.* Hartley and Sturm (1997); Kanatani et al. (2008); Lindstrom (2010). However, in some applications the nature of the scene is known, and thus, it imposes some extra restrictions to the 3D reconstruction. In many real-world scenes like the interior and facades of buildings Hoegner et al. (2016); Koch et al. (2019); Wefelscheid et al. (2011), city environments Poullis and You (2011); Sportouche et al. (2009); Laveau et al. (1998) or ground images from aerial vehicles (at high altitude) Caballero et al. (2009); Lu et al. (2018); Jurevičius and Marcinkevičius (2019), there exists a predominance of planar structures. Hence, being able to recover *and* guarantee that the reconstructed scene is planar is of special relevance for multiple and diverse applications.

Whereas the gold-standard approach (bundle-adjustment) addresses the *joint* problem of estimating both motion (pose displacement) and structure, see *e.g.* Faugeras and Lustman (1988); Bartoli and Sturm (2003); Bartoli et al. (2001), the non-convexity of the problem requires good initial guesses to converge to the optimal solution. On the other side, the relative motion can be decouple from the structure and estimated with high accuracy. Therefore, we focus here only on the triangulation problem given motion, which is still a non-convex problem with many local minima. Obtaining the 3D world point that originates the pair-wise correspondences is a trivial task in the noiseless case, since the "rays" emanating from the optical center towards these correspondences for each camera do intersect in the space at the real 3D point (see Figure (1)); its coordinates can be estimated, for example, with the linear method presented in Hartley and Zisserman (2003). When dealing with noisy correspondences, we can still leverage the same linear method to estimate a suboptimal solution. It is usually preferred in these cases to correct the observations so that the rays intersect in the space exactly. For 3D points on a plane (*planar triangula-*

**Corresponding author
e-mail: mercedesgarsal@uma.es (Mercedes Garcia-Salguero)

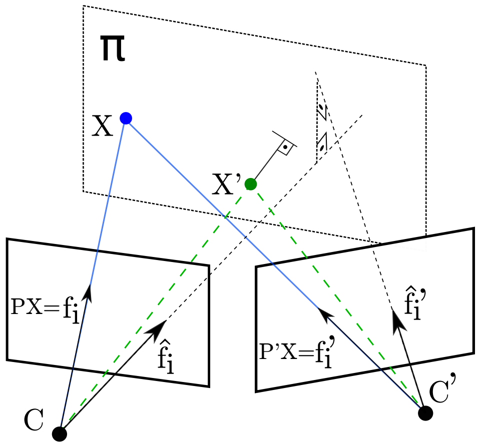


Fig. 1: The back-projection of noisy observations $(\hat{\boldsymbol{f}}, \hat{\boldsymbol{f}}')$ gives us rays that do not meet at a 3D point. To end up with an intersection we need to correct those observations in some optimal way such that a 3D point can be recovered (green point $\boldsymbol{X}'$). If the 3D point is known to lie on a plane π, some extra restrictions must be applied to the problem to obtain $\boldsymbol{X}$ (blue point), with corrected observations $(\boldsymbol{f}, \boldsymbol{f}')$

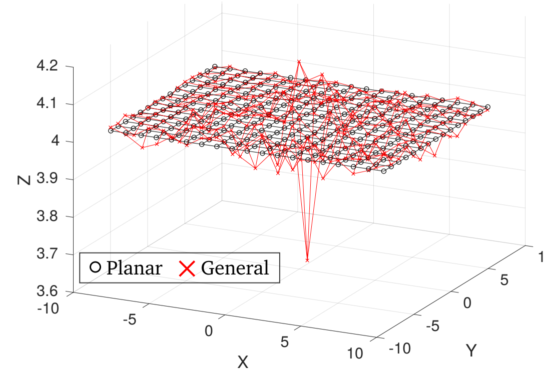


Fig. 2: Reconstructed points (red crosses) with a generic triangulation solver do not, in general, lie on the real 3D plane (black circles).

tion problem) Chum *et al*. Chum et al. (2005) proposed a polynomial solver, akin to the general one by Hartley and Sturm Hartley and Sturm (1997). The method requires finding all the roots of a 8-th degree polynomial, and later selecting the real one with the lower cost. Whereas this method is optimal, the tools required are usually unstable and/or slow. Kanazawa and Kanatani Kanazawa and Kanatani (1995) proposed earlier a first order approximation, which was later improved by Kanatani and Niitsuma in Kanatani and Niitsuma (2011). The last method, which tends to converge to the global optimum, is at its core iterative and thus, comes without optimality guarantees.

Contribution: In this work, we propose a set of certifiable solvers for the two-view planar triangulation problem. We base our proposals on duality theory Boyd and Vandenberghe (2004), which provides us with the necessary tools to certify optimal solutions to non-convex problems.

We start by proposing a certifiable solver in Algorithm (1) in Section (4) that both obtains a solution to the problem and certifies that such solution is the global optimum. Since other solvers may be employed to estimate the solution, for example, bundle-adjustment or any iterative solver akin to Kanatani et al. (2008), we present in Section (5) an optimality certificate to certify if a given solution to the problem is the global optimum. This certification algorithm, summarized in Alg. (2), is simple and only involves solving a linear least-squares system on two variables and a scalar computation, thus not requiring any specific optimization tool. Importantly, this certifier allows us to derive in Th. 5.2 an even simpler sufficient condition for optimality (not necessary) that eliminates the resolution of the linear least-squares system while being able to certify optimality in less than one microsecond.

The performance of the proposals is showcased with an extensive set of experiments on both synthetic and real data in Section (6). Therefore, our main results are provided in Alg. 1, Th. 5.1 (also Alg. 2) and Th. 5.2. We refer the reader that is only interested in their usage to these items and to Section 6 for the empirical evaluation of the proposals. We make the code available at `https://github.com/mergarsal`.

Notation: Through the paper, we denote matrices by bold upper-case letter, *e.g.* $\boldsymbol{A}, \boldsymbol{E}$, and by bold, lower-case letter, $\boldsymbol{b}$, (column) vectors. The symbol λ denotes the Lagrange multiplier. The non-negative orthant is $\mathbb{R}_+$, that is, $\mathbb{R}_+ \doteq \{a \in \mathbb{R} | a \geq 0\}$ and the set of $n \times n$ positive semidefinite (PSD) matrices by $\mathbb{S}^n_+$. For simplicity, we indicate that the matrix $\boldsymbol{H}$ is PSD as $\boldsymbol{H} \in \mathbb{S}^n_+$ or $\boldsymbol{H} \succeq 0$. The set of orthogonal matrices of dimension $n \times n$ is denoted by $\mathrm{O}(n) \doteq \{\boldsymbol{A} \in \mathbb{R}^{n \times n} | \boldsymbol{A}^T \boldsymbol{A} = \boldsymbol{I}_n, \boldsymbol{A}\boldsymbol{A}^T = \boldsymbol{I}_n\}$.

2. Related work

In this Section we summarize the most extended approaches to reconstruct 3D points on planes. We group the different approaches into two sets depending on which error they minimize in order to find the unknown 3D world point: (a) algebraic error; and (b) image error. To conclude the Section, we list certifiable algorithms similar to those proposed here that can be found for computer vision and robotics problems. We limit this Section to those works that consider two views and the triangulation problem. We refer the reader to Bartoli et al. (2001); Faugeras and Lustman (1988) and references therein for the structure and motion estimation from two images on planar scenes and Bartoli and Sturm (2003); Szeliski and Torr (1998) for an extension to multiple views.

Algebraic error: Given two noiseless observations from two cameras with known projective matrices, we can estimate the 3D point that originates them by the so-called linear method Hartley and Zisserman (2003); Hartley and Sturm (1997). When the observations are corrupted by noise, this method can still be leveraged and the solution is the one that minimizes an algebraic error. The solution in this case, though, may have a poor quality and other methods are usually preferred Hartley and Sturm (1997). An alternative solver is the so-called midpoint method Beardsley et al. (1994); Yang et al. (2019) that returns as solution the midpoint in the common perpendicular to the two rays, which is the point that minimizes the sum of squared distances to each ray. For the general case, Lee and Civera Lee and Civera (2019) proposed a weighted midpoint algorithm that was reported to recover good solutions close to the “optimal“ ones, which we list below.

All these solvers, though, do not consider the strong requirements of the points laying on a plane (see the comparison in Figure (2)). To overcome this, previous works see *e.g.* Micusik and Wildenauer (2017); Wefelscheid et al. (2011); Poullis and You (2011), fit the reconstructed (general) 3D points to the plane. The solution is only suboptimal and in the presence of noise, there is no guarantee that the fitted points are consistent with the plane derived from the homography matrix. Another approach that only requires one image, see *e.g.* Bartoli and Sturm (2001), reconstructs the 3D point from only one observation and the homography matrix by intersecting the ray from the observation with the 3D plane. Errors in the observation lead to an inconsistent structure for future images, that is, a corresponding (also noisy) observation on another image may not necessary be the image of the previously recovered 3D point. While the solutions to all these solvers may be fed to iterative algorithms for their refinement, there is no guarantee of convergence to the global optimum.

It is usually preferred to leverage another set of solvers defined as *optimal algorithms* Hartley and Zisserman (2003). Please notice that this was the original designated name, which in the context of this paper may lead to understandable confusion. In these *optimal algorithms* the term “optimal“ implies that the solvers correct the original observations so that they give rise to a unique point in the space, that is, the “rays“ from the camera centers towards the corrected observations intersect at a single point. Since the problem is non-convex, the solution does not need to be the global optimum if an iterative solver is used. The core of this paper relies on certifiable solvers, which we introduce later on this Section, that can actually certify optimality. When necessary, we will explicitly indicate if the solver is able to certify optimality or just correct the matches.

Image error: By assuming a pin-hole camera model, the image reprojection error is defined as the distance from the original observations to the image projection of the unknown 3D point. The corrected observations are sought to be the closest to the original ones in terms on some norm. For the general case the chosen norm, *e.g.* $\ell_2, \ell_1, \ell_\infty$, leads to different optimal solutions that attain different results regarding distance to the original 3D points and observations Lee and Civera (2019). For the planar triangulation problem, the different approaches have been limited to the minimization of the ℓ_2. This norm was shown in Chum et al. (2005) to be the Maximum Likelihood Estimation (MLE) under a Gaussian noise assumption. In the same paper, the authors proposed a polynomial solver, akin to the general one by Hartley and Sturm Hartley and Sturm (1997), that obtains the solution by computing the roots of a 8-th polynomial in one variable and selecting the real root with the lowest cost. However, polynomials solvers are slow and/or unstable, and more simple, although iterative approaches, are usually preferred. Kanazawa and Kanatani Kanazawa and Kanatani (1995) proposed earlier a first order approximation, which was later improved by Kanatani and Niitsuma in Kanatani and Niitsuma (2011). The last method, which tends to converge to the global optimum, is at its core iterative and thus, comes without optimality guarantees.

Nevertheless, there exist other approaches that both allow to obtain and certify optimal solutions for non-convex problems. Next we list the main approaches that stand on the same theory that the proposals on this manuscript, highlighting their advantages with respect to the other methods. We refer the interested reader to the references in those works for a complete comparison.

Certifiable algorithms: The two-view triangulation problem is non-convex with many local minima. A *tight* relaxation of the problem consists of a simpler problem that approximates the original one. Duality theory Boyd and Vandenberghe (2004) provides with a set of useful tools and relaxations that under some conditions (*strong duality*) allow to recover the solution to the original problem and certify it as optimal. Additionally, polynomial optimization based on Lassarre’s hierarchy provides with tighter although more complex relaxations. In this work, though, we focus on the dual problem and refer the reader to Lasserre (2008) for further information about the polynomial approach. Among the dual-based approaches, we must highlight one of these relaxations known as Shor’s relaxation, which for problems with quadratic objective and constraints (QCQP) can be solved by off-the-self tools with polynomial complexity. This approach is faster than the other family of global solvers that rely on Branch-and-Bound global optimization, with worst case exponential complexity. Shor’s relaxation appears in previous works, to name a few: pose graph optimization Rosen et al. (2019); Briales and Gonzalez-Jimenez (2017a), relative pose problem with central cameras Zhao (2020), generalized relative pose problem Zhao et al. (2020), the N-view triangulation problem Aholt et al. (2012); Cifuentes (2021) (the latter being a stronger albeit slower relaxation) and point cloud registration with outliers Yang and Carlone (2019). Another convex relaxation that appears in the literature is known as the dual problem. For QCQP, the dual problem can be also solved in polynomial time. Provided *strong duality* holds, this problem is able to certify solutions. From the solution to the dual problem and with strong duality, we can also estimate the solution to the original one if strict complementary also holds. These situations tend to be fulfilled for a wide set of problems and have been leveraged by previous works, *e.g.* SLAM Briales and Gonzalez-Jimenez (2016), registration with basic elements Briales and Gonzalez-Jimenez (2017b). Our first certifiable solver builds upon this last relaxation and jointly solved for the original (primal) and dual problems, allowing to both estimate and certify the returned solution as the global optimum.

Nonetheless, as it has been reported in the literature and is well-known by the research community, iterative solvers tend to converge to the global optimum of the non-convex problem, although without guarantees. Bearing in mind this behavior, previous works have proposed optimality certificates, that only certify solutions to the problem but not obtain them. Such candidates to optimal solutions can be estimated by other means in a efficient, although iterative way, thus the importance of the certification. These certificates are faster to compute than the above-mentioned solvers and for some problems, they consist on only a least-square system and a eigenvalue decomposition.

In Eriksson et al. (2018) the authors certify the rotation averaging problem. A further step was taken in Briales and Gonzalez-Jimenez (2016); Carlone et al. (2015); Carlone and Dellaert (2015) for the certification of solutions to the SLAM problem. A similar certification algorithm was presented in Iglesias et al. (2020) for the point cloud registration with missing data and Yang et al. (2020) for outliers. In Garcia-Salguero et al. (2021) we proposed the certification for the relative pose problem between calibrated cameras. The second part of this manuscript follows the same approach and propose the first optimality certification algorithm of this kind for the two-view planar triangulation problem.

Nevertheless, for some problems we can derive an even simpler optimality condition that is sufficient but not necessary. If this condition is *tight*, then it can certify most of the optimal solutions. The certification is usually faster than the certifier, thus making it a better option for applications in which speed of computation is a strong requirement. The published sufficient conditions are, however, very limited. In Eriksson et al. (2018) Eriksson *et al*. proposed such condition for the rotation averaging problem. This work was followed by Iglesias *et al*. in Iglesias et al. (2020) and extended to the point cloud registration including missing data. Hartley and Seo in Hartley and Seo (2008) bounded an optimality zone for the general N-view triangulation problem in which the cost was guaranteed to be convex. This bound depends on the cost of the solution, an upper and lower bound on the inverse depth of the 3D point and requires the computation of the eigenvalues of a 3×3 matrix. In contrast, the last part of this manuscript presents a sufficient condition for the two-view planar triangulation problem that directly bounds the norm of the optimal solution and whose slowest step consists in computing the least singular value of a 4×2 matrix.

3. Problem formulation

In this work we consider the two-view planar triangulation problem, where we assume the unknown 3D world point $\boldsymbol{X}$ belongs to a given plane π. Considering a perspective camera model Hartley and Zisserman (2003), this point $\boldsymbol{X}$ is projected onto the views via the projection matrices as the observations $\hat{\boldsymbol{f}}, \hat{\boldsymbol{f}}' \in \mathbb{R}^3$. Given the plane π, the cameras' projection matrices and the potentially noisy observations $\hat{\boldsymbol{f}}, \hat{\boldsymbol{f}}'$, we aim to reconstruct the 3D point $\boldsymbol{X}$ that originates them, by correcting these correspondences to $\boldsymbol{f}, \boldsymbol{f}' \in \mathbb{R}^3$, respectively, so that the rays from them meet at a 3D point that belongs to the plane π. Figure (2) illustrates why the second constraint cannot be omitted.

The plane π induces a homography $\boldsymbol{H}$ between the views Hartley and Sturm (1997), consequently relating the correspondences by $\boldsymbol{f}' \sim \boldsymbol{H}\boldsymbol{f}$, where $\sim$ denotes equality *up-to-scale*. To eliminate the scale ambiguity from this relation we take the cross-product as

$$[\boldsymbol{f}']_{\times}\boldsymbol{H}\boldsymbol{f} = \boldsymbol{0}_{3\times 1}. \tag{1}$$

Let us define the matrices

$$\boldsymbol{T}_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \ \boldsymbol{T}_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \ \boldsymbol{T}_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{2}$$

We can write the three equations in condition (1) as

$$\boldsymbol{f}'^T\boldsymbol{T}_1^T\boldsymbol{H}\boldsymbol{f} = 0, \ \boldsymbol{f}'^T\boldsymbol{T}_2^T\boldsymbol{H}\boldsymbol{f} = 0, \ \boldsymbol{f}'^T\boldsymbol{T}_3^T\boldsymbol{H}\boldsymbol{f} = 0. \tag{3}$$

Only two of the above three equations are algebraically independent, and thus two constraints[1] are necessary and sufficient for a pair of observations to be generated by a 3D point on the plane π.

We then seek the observations $(\boldsymbol{f}, \boldsymbol{f}')$ closest to the data $(\hat{\boldsymbol{f}}, \hat{\boldsymbol{f}}')$ such that (3) are fulfilled. Assuming Gaussian noise, the MLE is given when the distance function is the ℓ_2 Chum et al. (2005). We state the problem in terms of the corrections $\Delta\boldsymbol{f}, \Delta\boldsymbol{f}' \in \mathbb{R}^2$, although the formulation is equivalent to the one in Chum et al. (2005) and Kanatani and Niitsuma (2011):

$$\begin{aligned} f^\star = & \min_{\Delta f, \Delta f' \in \mathbb{R}^2 \times \mathbb{R}^2} \|\Delta\boldsymbol{f}\|_2 + \|\Delta\boldsymbol{f}'\|_2 \\ & \text{subject to } \begin{array}{l} (\hat{\boldsymbol{f}}' + \boldsymbol{G}^T\Delta\boldsymbol{f}')^T\boldsymbol{T}_1^T\boldsymbol{H}(\hat{\boldsymbol{f}} + \boldsymbol{G}^T\Delta\boldsymbol{f}) = 0 \\ (\hat{\boldsymbol{f}}' + \boldsymbol{G}^T\Delta\boldsymbol{f}')^T\boldsymbol{T}_2^T\boldsymbol{H}(\hat{\boldsymbol{f}} + \boldsymbol{G}^T\Delta\boldsymbol{f}) = 0 \end{array} \end{aligned} \tag{O}$$

where $\boldsymbol{G} = [\boldsymbol{I}_2 | \boldsymbol{0}_{2\times 1}] \in \mathbb{R}^{2\times 3}$. The variables of the problem are $\Delta\boldsymbol{f}, \Delta\boldsymbol{f}' \in \mathbb{R}^2$, which increments the first two entries of $\hat{\boldsymbol{f}}, \hat{\boldsymbol{f}}'$, respectively, and so $\boldsymbol{f} = \hat{\boldsymbol{f}} + \boldsymbol{G}^T\Delta\boldsymbol{f}$ (similar for $\boldsymbol{f}'$). Thus, we define the observations $\hat{\boldsymbol{f}}, \hat{\boldsymbol{f}}'$ on the normalized image plane, *i.e.*, they are homogeneous 3D vectors with last entry one. Thus, the corrected observations $\boldsymbol{f}, \boldsymbol{f}'$ also belong to the normalized image plane. Notice that the error minimized in this problem is indeed geometric Chum et al. (2005).

This problem, though, it is non-convex and NP-hard to solve. We can, nevertheless, leverage concepts from duality theory in order to solve and certificate solutions in a simpler manner. For that, we need to re-write problem (O) in a more standard form. We define the vector of unknowns $\boldsymbol{w} = [\Delta\boldsymbol{f}^T, \Delta\boldsymbol{f}'^T]^T \in \mathbb{R}^4$ and $\boldsymbol{x} = [\boldsymbol{w}^T, y]^T \in \mathbb{R}^5$, with $y \in \mathbb{R}$ the element (and later variable) that makes the constraints homogeneous. In terms of $\boldsymbol{x}$ we can re-write the constraints as $(\hat{\boldsymbol{f}}' + \boldsymbol{G}^T\Delta\boldsymbol{f}')^T\boldsymbol{T}_1^T\boldsymbol{H}(\hat{\boldsymbol{f}} + \boldsymbol{G}^T\Delta\boldsymbol{f}) = \boldsymbol{w}^T\boldsymbol{A}_1\boldsymbol{w} + y2\boldsymbol{b}_1^T\boldsymbol{w} + c_1y^2 = 0 \Leftrightarrow \boldsymbol{x}^T\boldsymbol{B}_1\boldsymbol{x} = 0$ (and similar for the second constraint) with

$$\boldsymbol{B}_i = \begin{pmatrix} \boldsymbol{A}_i & \boldsymbol{b}_i \\ \boldsymbol{b}_i^T & c_i \end{pmatrix} \in \mathbb{S}^5, \ i = 1, 2 \tag{4}$$

with $\mathbb{S}^5$ the space of symmetric matrices of size 5. Due to the homogenization we incorporate the additional constraint $y^2 = 1 \Leftrightarrow \boldsymbol{x}^T\boldsymbol{L}\boldsymbol{x} = 1$ with $\boldsymbol{L} = \boldsymbol{0}_{4\times 4} \oplus 1 \in \mathbb{S}^5_+$, where $\oplus$ is the direct sum of two matrices and $\mathbb{S}^5_+$ the cone of positive semidefinite matrices of size 5. We write the cost in terms of the variable $\boldsymbol{x}$ as $\boldsymbol{w}^T\boldsymbol{w} = \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$ with $\boldsymbol{Q} = \boldsymbol{I}_4 \oplus 0 \in \mathbb{S}^5_+$.

The standard problem is given by

$$\begin{aligned} f^\star = & \min_{\boldsymbol{x} \in \mathbb{R}^5} \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} \\ & \text{subject to } \boldsymbol{x}^T\boldsymbol{B}_1\boldsymbol{x} = 0, \quad \boldsymbol{x}^T\boldsymbol{B}_2\boldsymbol{x} = 0, \quad \boldsymbol{x}^T\boldsymbol{L}\boldsymbol{x} = 1 \end{aligned} \tag{QCQP}$$

[1] We use the first two: $\boldsymbol{f}'^T\boldsymbol{T}_1^T\boldsymbol{H}\boldsymbol{f} = 0$ and $\boldsymbol{f}'^T\boldsymbol{T}_2^T\boldsymbol{H}\boldsymbol{f} = 0$.

While this problem is still non-convex, it allows us to derive a convex relaxation, the so-called *dual problem* Boyd and Vandenberghe (2004), that can certify and even obtain the optimal solution provided *strong duality* holds. We state next the main results from this convexification and refer the reader to Appendix A for the full derivation of the dual problem and the set of conditions. Following the standard procedure, see (Boyd and Vandenberghe, 2004, App. B.1), the dual problem has the final form:

$$g^\star = \max_{\lambda_1, \lambda_2, \rho \in \mathbb{R}} \rho \tag{D}$$
$$\text{subject to } \boldsymbol{H}_\lambda \doteq \begin{pmatrix} \boldsymbol{I}_4 - \lambda_1 \boldsymbol{A}_1 - \lambda_2 \boldsymbol{A}_2 & -(\lambda_1 \boldsymbol{b}_1 + \lambda_2 \boldsymbol{b}_2) \\ -(\lambda_1 \boldsymbol{b}_1 + \lambda_2 \boldsymbol{b}_2)^T & \rho - \lambda_1 c_1 - \lambda_2 c_2 \end{pmatrix} \succeq 0$$

where the symbol $\succ$ (resp. $\succeq$) implies positive definite PD (resp. positive semidefinite PSD), and $\boldsymbol{H}_\lambda$ denotes the Hessian of the Lagrangian.

Necessary and sufficient optimality conditions Assuming strong duality holds $g^\star = f^\star$, we can derive the next sufficient and necessary conditions. Note that if strong duality doesn't hold, these conditions don't provide any information since we have that $g^\star < f^\star$ with strict inequality. We seek points for the primal $\boldsymbol{x} = [\boldsymbol{w}^T, 1]$ and dual problem $(\lambda_1, \lambda_2, \rho)$ such that the following conditions are fulfilled:

$$\boldsymbol{I}_4 - \lambda_1 \boldsymbol{A}_1 - \lambda_2 \boldsymbol{A}_2 \succeq 0 \tag{5}$$
$$(\boldsymbol{I}_4 - \lambda_1 \boldsymbol{A}_1 - \lambda_2 \boldsymbol{A}_2)\boldsymbol{w} = (\lambda_1 \boldsymbol{b}_1 + \lambda_2 \boldsymbol{b}_2) \tag{6}$$
$$\rho = \boldsymbol{w}^T \boldsymbol{w} \tag{7}$$
$$\boldsymbol{w}^T \boldsymbol{A}_1 \boldsymbol{w} + 2\boldsymbol{b}_1^T \boldsymbol{w} + c_1 = 0 \tag{8}$$
$$\boldsymbol{w}^T \boldsymbol{A}_2 \boldsymbol{w} + 2\boldsymbol{b}_2^T \boldsymbol{w} + c_2 = 0 \tag{9}$$

An alternative convex relaxation, known as the Shor's relaxation, that also allows to obtain and certify the optimal solution under some conditions is provided in the *Supplementary material* Section (Appendix B).

Solution to system (5)-(9) Therefore, we seek the solution(s) to the system of equations in (5)-(9). From Equation (6), we have that $\boldsymbol{w} = \boldsymbol{S}_\lambda^{-1}(\lambda_1 \boldsymbol{b}_1 + \lambda_2 \boldsymbol{b}_2)$, which is unique if $\boldsymbol{S}_\lambda$ has full-rank. In what follows, we assume that the matrix *is* positive definite. While this seems a strong assumption, we observe experimentally in Section (6) that it is the case in all the experiments. A proof similar to that in Hmam (2010) is being considered but let as future work. Assuming positive definiteness, we can reformulate the constraints in (8) and (9) in terms of λ_1, λ_2, as polynomials of degree 9. From all the solutions to this polynomial system of two equations in two variables, we are interested in those that make $\boldsymbol{S}_\lambda \succ 0$. The set of conditions in Eqs. (5)-(9) is leveraged by the proposed certifiable algorithms in Sections 4 and 5. Note that for the problem at hand, this set has a special form since the data follows a the pattern given by Prob. QCQP.

4. Efficient primal-dual solver

First, we propose a fast primal-dual solver that (1) estimates the solution of Prob. QCQP and (2) certifies it as the global optimum. This solver relies on the set of Eqs. (5)-(9), hence assuming strong duality. Further, given the pattern of the data matrices, this solver can be efficiently implemented, as we explain in this Section. The primal-dual solver is summarized in Alg. 1, whereas the remaining of this Section focuses on the development of the expressions for the solver.

Finding the solution to Eq. (A.9) involves the inversion of the Hessian $\boldsymbol{H}_\lambda$, which is evaluated at each potential candidate solution (λ_1, λ_2). We propose here a re-formulation of the problem that avoids the computation of the inverse, thus simplifying the form of the constraints and the condition that $\boldsymbol{S}_\lambda \succ 0$. The transformation is similar to that proposed in Hmam (2010). For the considered constraints in problem (QCQP) we have that

$$\boldsymbol{G}\boldsymbol{T}_1^T \boldsymbol{H} \boldsymbol{G}^T = \begin{pmatrix} 0 & 0 \\ h_{3,1} & h_{3,2} \end{pmatrix}, \ \boldsymbol{G}\boldsymbol{T}_2^T \boldsymbol{H} \boldsymbol{G}^T = \begin{pmatrix} -h_{3,1} & -h_{3,2} \\ 0 & 0 \end{pmatrix}, \tag{10}$$

where $h_{i,j}$ is the (i, j)-th entry of $\boldsymbol{H}$ and $\boldsymbol{T}_1, \boldsymbol{T}_2$ the 3×3 matrices given in (2). Let us consider the SVD of $\boldsymbol{T}_2^T \boldsymbol{H} = \boldsymbol{U}\boldsymbol{D}\boldsymbol{V}^T$, with $\boldsymbol{D} = \text{diag}(s, 0)$ and $s = \frac{1}{2}(h_{3,1}^2 + h_{3,2}^2)^{1/2}$. Due to the structure of $\boldsymbol{T}_1^T \boldsymbol{H}$, we also have that

$$\boldsymbol{U}^T \boldsymbol{G}\boldsymbol{T}_1^T \boldsymbol{H} \boldsymbol{G}^T \boldsymbol{V} = \begin{pmatrix} 0 & 0 \\ -s & 0 \end{pmatrix}. \tag{11}$$

Let us define the orthogonal matrix

$$\boldsymbol{O} = \begin{pmatrix} \boldsymbol{0}_{2\times2} & \boldsymbol{U} \\ \boldsymbol{V}^T & \boldsymbol{0}_{2\times2} \end{pmatrix}, \tag{12}$$

and the vector $\boldsymbol{v} = \boldsymbol{O}\boldsymbol{w} \in \mathbb{R}^4$. We can write the original problem (QCQP) in terms of $\boldsymbol{v}$ by a change of variables. The cost function $f(\boldsymbol{w}) = \boldsymbol{w}^T \boldsymbol{w} = \boldsymbol{v}^T \boldsymbol{v}$ has the same value under the change for any point. For the constraints we define the matrices $\boldsymbol{C}_1 = \boldsymbol{O}\boldsymbol{A}_1 \boldsymbol{O}^T$ and the vector $\boldsymbol{r}_1 = \boldsymbol{O}\boldsymbol{b}_1 \in \mathbb{R}^4$ (similar for the second expression) and $\boldsymbol{w}^T \boldsymbol{A}_1 \boldsymbol{w} + 2\boldsymbol{b}_1^T \boldsymbol{w} + c_i = \boldsymbol{v}^T \boldsymbol{C}_1 \boldsymbol{v} + 2\boldsymbol{r}_1^T \boldsymbol{v} + c_i = 0$. Following the same procedure, we obtain optimality conditions of the form of (5)-(9) but in terms of $\boldsymbol{v}, \boldsymbol{C}_i, \boldsymbol{r}_i$. Nonetheless, for this representation we have that

$$\boldsymbol{I}_4 - \lambda_1 \boldsymbol{C}_1 - \lambda_2 \boldsymbol{C}_2 = \begin{pmatrix} 1 & 0 & -\lambda_2 s & 0 \\ 0 & 1 & \lambda_1 s & 0 \\ -\lambda_2 s & \lambda_1 s & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{13}$$

whose inverse is given by

$$(\boldsymbol{I}_4 - \lambda_1 \boldsymbol{C}_1 - \lambda_2 \boldsymbol{C}_2)^{-1} = \frac{1}{s^2(\lambda_1^2 + \lambda_2^2) - 1} \begin{pmatrix} s^2\lambda_1^2 - 1 & s^2\lambda_1\lambda_2 & -s\lambda_2 & 0 \\ s^2\lambda_1\lambda_2 & s^2\lambda_2^2 - 1 & s\lambda_1 & 0 \\ -s\lambda_2 & s\lambda_1 & -1 & 0 \\ 0 & 0 & 0 & s^2(\lambda_1^2 + \lambda_2^2) - 1 \end{pmatrix} \tag{14}$$

With a little abuse of notation, let us denoted by $\boldsymbol{S}_\lambda$ the new matrix $\boldsymbol{S}_\lambda = \boldsymbol{I}_4 - \lambda_1 \boldsymbol{C}_1 - \lambda_2 \boldsymbol{C}_2$. The sum $\lambda_1 \boldsymbol{C}_1 + \lambda_2 \boldsymbol{C}_2$ has as eigenvalues $\{-m, m, 0, 0\}$ with $m = \frac{1}{2}(h_{3,1}^2 + h_{3,2}^2)^{1/2}(\lambda_2^2 + \lambda_1^2)^{1/2} = s(\lambda_2^2 + \lambda_1^2)^{1/2}$. For $\boldsymbol{S}_\lambda$ to be PSD, it is required that

$$1 - m \geq 0 \Leftrightarrow 2/(h_{3,1}^2 + h_{3,2}^2)^{1/2} \geq (\lambda_2^2 + \lambda_1^2)^{1/2}. \tag{15}$$

Algorithm 1: Primal-dual solver for two-view planar triangulation

Data: Noisy observations $(\hat{\boldsymbol{f}}, \hat{\boldsymbol{f}}')$, homography matrix $\boldsymbol{H}$
Result: Corrected observations $(\boldsymbol{f}, \boldsymbol{f}')$; optimality certificate isOpt

```
// Compute transformation
Compute SVD of GT_2^T HG^T: set s, O;
// Compute data for observation
Compute vectors r_1 = O b_1, r_2 = O b_2;
Compute scalars c_1 = f'^T T_1^T H f, c_2 = f'^T T_2^T H f ;
Compute initial guess, e.g. λ_10 = -1/2 c_1/(r_1^T r_1), λ_20 = -1/2 c_2/(r_2^T r_2)
Compute coefficients for Newton's method[2]
// Find zero
repeat
    Evaluate constraints k_1(λ_1, λ_2) in Eq. (16) and k_2(λ_1, λ_2) in Eq. (17) ;
    Evaluate Jacobian (J(λ_1i, λ_2i))^{-1} (closed-form);
    Update multipliers with Eq. (18);
until convergence or max. iters;
Estimate v* = (I_4 - λ_1* C_1 - λ_2* C_2)^{-1} (λ_1* r_1 + λ_2* r_2) (Eq. (14));
Transform back w* = O^T v*;
Update observations with w*;
if λ_1*^2 + λ_2*^2 < (1 - ε_min)/s^2 then
    // Solution is optimal
    isOpt = True ;
else
    isOpt = unknown ;
```

We then need to find the solution to the polynomial system

$$
\begin{aligned}
k_1(\lambda_1, \lambda_2) &= (\lambda_1 \boldsymbol{r}_1 + \lambda_2 \boldsymbol{r}_2)^T \boldsymbol{S}_\lambda^{-T} \boldsymbol{C}_1 \boldsymbol{S}_\lambda^{-1} (\lambda_1 \boldsymbol{r}_1 + \lambda_2 \boldsymbol{r}_2) \\
&\quad + 2 \boldsymbol{r}_1^T \boldsymbol{S}_\lambda^{-1} (\lambda_1 \boldsymbol{r}_1 + \lambda_2 \boldsymbol{r}_2) + c_1 = 0 \qquad (16) \\
k_2(\lambda_1, \lambda_2) &= (\lambda_1 \boldsymbol{r}_1 + \lambda_2 \boldsymbol{r}_2)^T \boldsymbol{S}_\lambda^{-T} \boldsymbol{C}_2 \boldsymbol{S}_\lambda^{-1} (\lambda_1 \boldsymbol{r}_1 + \lambda_2 \boldsymbol{r}_2) \\
&\quad + 2 \boldsymbol{r}_2^T \boldsymbol{S}_\lambda^{-1} (\lambda_1 \boldsymbol{r}_1 + \lambda_2 \boldsymbol{r}_2) + c_2 = 0 \qquad (17)
\end{aligned}
$$

such that $2/(h_{3,1}^2 + h_{3,2}^2)^{1/2} \geq (\lambda_2^2 + \lambda_1^2)^{1/2}$.

Among all the methods that can be employed to solve this system, we choose Newton's method, which requires of a good initialization to converge to the solution. By noting that the first term of the constraints tends to zero and the matrix $\boldsymbol{S}_\lambda$ to the identity of size four, we propose to initialize each variable by $\lambda_{10} = -\frac{1}{2}\frac{c_1}{\boldsymbol{r}_1^T \boldsymbol{r}_1}$ and $\lambda_{20} = -\frac{1}{2}\frac{c_2}{\boldsymbol{r}_2^T \boldsymbol{r}_2}$.

Last, Newton's method updates the variables at i-th iteration as

$$
\begin{pmatrix} \lambda_{1i+1} \\ \lambda_{2i+1} \end{pmatrix} = \begin{pmatrix} \lambda_{1i} \\ \lambda_{2i} \end{pmatrix} - (\boldsymbol{J}(\lambda_{1i}, \lambda_{2i}))^{-1} \begin{pmatrix} k_1(\lambda_{1i}, \lambda_{2i}) \\ k_2(\lambda_{1i}, \lambda_{2i}) \end{pmatrix} \qquad (18)
$$

being $\boldsymbol{J}(\lambda_{1i}, \lambda_{2i}) \in \mathbb{R}^{2\times 2}$ the Jacobian of $[k_1, k_2]$ w.r.t. each variable whose entries are polynomials in λ_1, λ_2. The inverse has an explicit (and simple) form since the matrix is 2×2. The coefficients for both the Jacobian and k_1, k_2 depend only on the problem data and then, they are computed just once for problem instance.

With these, we can finally provide the algorithm in Alg. (1) that both estimates the solution to (QCQP) and an optimality certificate for it. We set the convergence when the increase between consecutive estimations remains below a threshold $\epsilon_{update} = 5e - 17$ or the value of the constraints drops to $\epsilon_{constraints} = 5e-17$. We apply a threshold $\epsilon_{min} = +1e-05$ to the PSD condition for the Hessian $\boldsymbol{H}_\lambda$. Please, note that all these thresholds were employed during our evaluation in Section 6. However, they can be adjusted by the users and we provide these values here for simplicity and as guidelines. For safety reasons, we limit the number of iterations to 10, although we observe in Section (6) an average of $3-4$ iterations even with highly noisy observations. The highlighted part marks the variables that can be re-used for all the problem instances under the same camera projections $\boldsymbol{P}, \boldsymbol{P}'$.

Last, let us remark that currently our solver does not guarantee to find the optimal solution since a pair of scalars (λ_1, λ_2) fulfilling constraints k_1, k_2 may not exist. We observe experimentally in Section (6) that a feasible dual point (λ_1, λ_2) does exist in all the tested problem instances. A theoretical proof of this behavior is projected as future work.

Affine case: When the homography is affine, $h_{3,1} = h_{3,2} = 0$ and so $s = 0$, then the PSD condition in (5) is trivially satisfied with $\boldsymbol{S}_\lambda = \boldsymbol{I}_4 \succ 0$. Therefore, the norm of the lagrange multipliers is bounded above by $2/(h_{3,1}^2 + h_{3,2}^2)^{1/2} = 2/0$ (see also Eq. (15)), that is, any solution to the system in Eqs. (16), (17) is the global optimum. The solution is estimated in closed-form, as pointed out in Chum et al. (2005). In this case, the optimal multipliers are computed as

$$
\begin{pmatrix} \lambda_1^\star \\ \lambda_2^\star \end{pmatrix} = \frac{1}{2} \frac{1}{\boldsymbol{r}_1^T \boldsymbol{r}_1 \boldsymbol{r}_2^T \boldsymbol{r}_2 - \boldsymbol{r}_2^T \boldsymbol{r}_1 \boldsymbol{r}_1^T \boldsymbol{r}_2} \begin{pmatrix} \boldsymbol{r}_2^T \boldsymbol{r}_2 & -\boldsymbol{r}_2^T \boldsymbol{r}_1 \\ -\boldsymbol{r}_1^T \boldsymbol{r}_2 & \boldsymbol{r}_1^T \boldsymbol{r}_1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \qquad (19)
$$

and the primal solution is $\boldsymbol{w}^\star = \lambda_1^\star \boldsymbol{r}_1 + \lambda_2^\star \boldsymbol{r}_2$. Please, note that the initialization presented above follows from a simplification of this case.

5. Optimality certificates

Other methods can be employed to estimate a solution to Problem (O) (not necessarily the optimum), see Section (1). It is useful then to be able to *certify* if the given solution is the global optimum. This section shows how to compute this optimality certificate and then to derive a simpler sufficient condition for optimality. As the primal-dual solver, the next two conditions are derived from Eqs. (5)-(9) and assume strong duality holds. Further, the results of this Section, collected in Th. 5.1 and Th. 5.2, also assume a primal solution is given, *i.e.* $\boldsymbol{f}, \boldsymbol{f}'$ and equivalently $\boldsymbol{x}$, $\boldsymbol{v}$, $\boldsymbol{w}$. Following the end results on these theorems, we provide their derivation which is not required to employ them.

Optimality certificate: The optimality certificate is based on the optimality conditions given in Eqs. (5)-(9) and is summarized in the next Theorem.

[2] The coefficients for the constraints and the entries of the Jacobian can be found in plain format in https://github.com/mergarsal

Algorithm 2: Optimality certificate

Data: Corrected observations $(\boldsymbol{f}, \boldsymbol{f}')$, homography matrix $\boldsymbol{H}$

Result: optimality certificate isOpt

```
// Compute candidate to multiplier
Compute λ̂ from Eq. (20);
// Compute minimum eigenvalue S_λ
Compute scalar m = ½(h²_{3,1} + h²_{3,2})^{1/2}(λ²_2 + λ²_1)^{1/2};
// Check dual gap (sanity)
dual gap = w^T w + λ1 c1 + λ2 c2 + (λ1 b1 + λ2 b2)^T w
if dual gap < ε_gap and 1 − m > +ε_min then
    // Solution is optimal
    isOpt = True ;
else
    isOpt = unknown ;
```

Theorem 5.1 (Optimality certificate). *Let $\boldsymbol{w} = (\Delta\boldsymbol{f}, \Delta\boldsymbol{f}')$ be a potentially optimal solution for the primal problem in (QCQP). The entries of the 2D vector $\boldsymbol{\lambda} = [\lambda_1, \lambda_2]^T$ are computed as*

$$\underbrace{(\boldsymbol{A}_1\boldsymbol{w} + \boldsymbol{b}_1 | \boldsymbol{A}_2\boldsymbol{w} + \boldsymbol{b}_2)}_{\mathbb{R}^{4\times 2}}\,\boldsymbol{\lambda} = \underbrace{\boldsymbol{w}}_{\mathbb{R}^4}\,. \tag{20}$$

and are the candidates to Lagrange multipliers. If the Hessian evaluated at this candidate $\boldsymbol{\lambda}$ as $\boldsymbol{S}_\lambda = \boldsymbol{I}_4 - \lambda_1\boldsymbol{A}_1 - \lambda_2\boldsymbol{A}_2$ is PSD and $\boldsymbol{w}^T\boldsymbol{w} = -(\lambda_1\boldsymbol{b}_1 + \lambda_2\boldsymbol{b}_2)^T\boldsymbol{S}_\lambda^{-1}(\lambda_1\boldsymbol{b}_1 + \lambda_2\boldsymbol{b}_2) - \lambda_1 c_1 - \lambda_2 c_2 = \rho$ holds, then it follows that (1) strong duality holds between (QCQP) and (D); (2) the solution $\boldsymbol{w}$ is the optimum of (QCQP); and (3) the solution $(\lambda_1, \lambda_2, \rho)$ is the optimum to (D).

Proof 5.1. *The primal and dual solutions fulfill the set of constraints in the system (5)-(9).* □

Algorithm (2) summarizes the optimality certificate derived from Th. (5.1). Due to numerical errors, we apply two thresholds to the optimality conditions; (1) feasibility, *i.e.* the matrix $\boldsymbol{S}_\lambda$ being PSD, is substituted to $\boldsymbol{S}_\lambda > +\epsilon_{min}\boldsymbol{I}_4$ with $\epsilon_{min} = 1e-5$; and (2) strong duality (dual gap equals zero) is defined as dual gap $\equiv |f(\boldsymbol{w}) - d(\lambda_1, \lambda_2)| \leq \epsilon_{gap} = 1e-10$. As in Alg. 1, these thresholds are included as guidelines and the users are allowed to change them if required. Note that the Alg. (2) can be inconclusive about the optimality of the given solution. This may happen if the solution is not optimal or if the underlying dual problem is not a tight relaxation of the problem (QCQP). In our extensive experiments, this situation was not found.

Sufficient condition for optimality: While Theorem (5.1) is already simple, it allows us to derive a more straightforward certificate without explicitly computing the lagrange multipliers $\boldsymbol{\lambda}$. The sufficient condition for optimality is summarized in the next Theorem

Theorem 5.2 (Sufficient condition). *For a problem instance with the form in Prob. O, with homography matrix $\boldsymbol{H} = \{h_{i,j}\}_{i,j=1}^3 \in \mathbb{R}^3$ and data $\boldsymbol{B} = [\boldsymbol{b}_1, \boldsymbol{b}_2] \in \mathbb{R}^{4\times 2}$, $\boldsymbol{b}_1 = \frac{1}{2}[\boldsymbol{f}'^T\boldsymbol{T}_1^T\boldsymbol{H}\boldsymbol{G}^T, \boldsymbol{f}^T\boldsymbol{H}^T\boldsymbol{T}_1\boldsymbol{G}^T]^T \in \mathbb{R}^4$ (similar for $\boldsymbol{b}_2$) any feasible primal solution $\boldsymbol{w} \in \mathbb{R}^4$ with norm bounded above by*

$$\|\boldsymbol{w}\|_2 \leq (1 - \epsilon_{min})\sigma(\boldsymbol{B})/(h_{3,1}^2 + h_{3,2}^2)^{1/2} \tag{21}$$

for a threshold $\epsilon_{min} > 0$ is the global optimum for the problem, which follows from the dual problem be tight.

Proof 5.2. *Let us employ again the transformed problem with data $\boldsymbol{C}_i, \boldsymbol{r}_i, c_i$ and variable $\boldsymbol{v} \in \mathbb{R}^4$. It was shown that the Hessian $\boldsymbol{S}_\lambda$ is PSD* iff $(1-\epsilon_{min})2/(h_{3,1}^2+h_{3,2}^2)^{1/2} \geq (\lambda_2^2+\lambda_1^2)^{1/2} = \|\boldsymbol{\lambda}\|_2$, *where we have introduced once again the threshold ϵ_{min}. Since, in the least-squares sense, the 2D vector $\boldsymbol{\lambda} = (\boldsymbol{C}_1\boldsymbol{v} + \boldsymbol{r}_1 | \boldsymbol{C}_2\boldsymbol{v} + \boldsymbol{r}_2)^\dagger\boldsymbol{v}$, we can bound from above the norm of $\boldsymbol{\lambda}$ and so*

$$\|\boldsymbol{\lambda}\|_2 \leq \left\|(\boldsymbol{C}_1\boldsymbol{v} + \boldsymbol{r}_1 | \boldsymbol{C}_2\boldsymbol{v} + \boldsymbol{r}_2)^\dagger\right\|_2 \|\boldsymbol{v}\|_2 \tag{22}$$

$$\leq 2(1 - \epsilon_{min})/(h_{3,1}^2 + h_{3,2}^2)^{1/2}, \tag{23}$$

where $\left\|(\boldsymbol{C}_1\boldsymbol{v} + \boldsymbol{r}_1 | \boldsymbol{C}_2\boldsymbol{v} + \boldsymbol{r}_2)^\dagger\right\|_2$ denotes the 2-norm of the pseudo-inverse of the matrix $\boldsymbol{C}_v \doteq (\boldsymbol{C}_1\boldsymbol{v} + \boldsymbol{r}_1 | \boldsymbol{C}_2\boldsymbol{v} + \boldsymbol{r}_2) \in \mathbb{R}^{4\times 2}$. In this case, the norm is equal to $1/\sigma$ with $\sigma = \sigma(\boldsymbol{C}_v)$ the minimum singular value of $\boldsymbol{C}_v$. Please, observe that $1/\sigma\|\boldsymbol{v}\|_2 \leq (1 - \epsilon_{min})2/(h_{3,1}^2 + h_{3,2}^2)^{1/2}$ is sufficient *for $\boldsymbol{S}_\lambda$ to be PSD but not necessary. Now, let us re-write the matrix $\boldsymbol{C}_v$ as the sum $\boldsymbol{C}_v = (\boldsymbol{C}_1\boldsymbol{v}|\boldsymbol{C}_2\boldsymbol{v}) + (\boldsymbol{r}_1|\boldsymbol{r}_2)$ and for short, $\boldsymbol{B} = (\boldsymbol{r}_1|\boldsymbol{r}_2) \in \mathbb{R}^{4\times 2}$. Applying Weyl's inequality to the minimum singular value $\sigma(\boldsymbol{C}_v)$ we have that:*

$$|\sigma(\boldsymbol{C}_v) - \sigma(\boldsymbol{B})| \leq \|(\boldsymbol{C}_1\boldsymbol{v}|\boldsymbol{C}_2\boldsymbol{v})\|_2. \tag{24}$$

Our objective is to bound $\sigma(\boldsymbol{C}_v)$ from below. Assume that $\sigma(\boldsymbol{B}) \geq \sigma(\boldsymbol{C}_v)$ then

$$|\sigma(\boldsymbol{C}_v) - \sigma(\boldsymbol{B})| = \sigma(\boldsymbol{B}) - \sigma(\boldsymbol{C}_v) \leq \|(\boldsymbol{C}_1\boldsymbol{v}|\boldsymbol{C}_2\boldsymbol{v})\|_2 \Leftrightarrow$$

$$\Leftrightarrow \sigma(\boldsymbol{B}) - \|(\boldsymbol{C}_1\boldsymbol{v}|\boldsymbol{C}_2\boldsymbol{v})\|_2 \leq \sigma(\boldsymbol{C}_v). \tag{25}$$

In the other case in which $\sigma(\boldsymbol{B}) \leq \sigma(\boldsymbol{C}_v)$ we have that $\sigma(\boldsymbol{B}) - \|(\boldsymbol{C}_1\boldsymbol{v}|\boldsymbol{C}_2\boldsymbol{v})\|_2 \leq \sigma(\boldsymbol{C}_v)$ still holds since the norm-2 of any matrix is always non-negative. We then have that

$$\|\boldsymbol{w}\|_2 \leq 2(1 - \epsilon_{min})/(h_{3,1}^2 + h_{3,2}^2)^{1/2}(\sigma(\boldsymbol{B}) - \|(\boldsymbol{C}_1\boldsymbol{v}|\boldsymbol{C}_2\boldsymbol{v})\|_2)$$

$$\leq 2(1 - \epsilon_{min})/(h_{3,1}^2 + h_{3,2}^2)^{1/2}\sigma(\boldsymbol{C}_v) \tag{26}$$

since $(1-\epsilon_{min})2/(h_{3,1}^2+h_{3,2}^2)^{1/2}$ is non-negative by construction. The structure of $(\boldsymbol{C}_1\boldsymbol{v}|\boldsymbol{C}_2\boldsymbol{v})$ is

$$(\boldsymbol{C}_1\boldsymbol{v}|\boldsymbol{C}_2\boldsymbol{v}) = \frac{1}{2}(h_{3,1}^2 + h_{3,2}^2)^{1/2}\begin{pmatrix} 0 & v_3 \\ -v_3 & 0 \\ -v_2 & v_1 \\ 0 & 0 \end{pmatrix}, \tag{27}$$

and so its singular values are given by $1/2(h_{3,1}^2 + h_{3,2}^2)^{1/2}(v_3^2)^{1/2}$ and $1/2(h_{3,1}^2 + h_{3,2}^2)^{1/2}(v_1^2 + v_2^2 + v_3^2)^{1/2}$. The norm-2 of the matrix is the maximum singular value and we can re-formulate the Equation in (26) as

$$\|\boldsymbol{v}\|_2 + \|(\boldsymbol{I}_3 \oplus 0)\boldsymbol{v}\|_2 \leq 2\|\boldsymbol{v}\|_2 \tag{28}$$

$$\leq 2(1 - \epsilon_{min})/(h_{3,1}^2 + h_{3,2}^2)^{1/2}\sigma(\boldsymbol{B})$$

from which it follows that $\|\boldsymbol{w}\|_2 = \|\boldsymbol{v}\|_2 \leq (1-\epsilon_{min})\sigma(\boldsymbol{B})/(h_{3,1}^2 + h_{3,2}^2)^{1/2}$*. Observe that the RHS of the above equation does not depend of the solution and that the term* $(1-\epsilon_{min})\sigma(\boldsymbol{B})/(h_{3,1}^2 + h_{3,2}^2)^{1/2}$ *provides an upper bound on the cost of the problem, that is, all the feasible solutions whose costs remains below this value are guaranteed to be optimal.*

Last, notice that for the affine case the norm of the solution is bounded above by $\sigma(\boldsymbol{B})/0$, that is, any feasible solution is the global optimum (see Section (4)). Next section is devoted to showcase that the proposed solvers and certificates do work in practice.

6. Evaluation

In this section we evaluate experimentally the performance of the solver in Algorithm (1), denoted as DUAL, and the certificates in Alg (2) and the condition in Eq. (21) on both synthetic (in Section (6.1)) and real data (in Section (6.2)). We compare the solvers against the polynomial solver (named as POLY) in Chum et al. (2005) that also certifies optimality and against the iterative solver in Kanatani and Niitsuma (2011) (denoted as ITER), which does not guarantee that the returned solution is the global optimum. We also include the polynomial solver for the general triangulation Hartley and Sturm (1997) HS-G and the solution returned by the linear method SVD Hartley and Zisserman (2003). Notice that HS-G and POLY follow the same approach and the main difference between them is that POLY is devised for planar scenes. We implement all the solvers according to the original papers and for the polynomial resolution, we employ the algorithm provided in `http://www.crbond.com/download/misc/rpoly.cpp`.

6.1. Evaluation on synthetic data

We generate a grid (see Figure (4)) of size 9×7 units with 15 columns and 19 rows (285 evenly-space points) on a plane with normal $[0,0,1]$ w.r.t. the world frame $\boldsymbol{W}$ and distance from the origin d_0 units (Z-axis). Each point on this grid is estimated separately since the considered algorithms only admit two views. We place the first camera $\boldsymbol{C}$ at a random pose with maximum angle of rotation 0.5 rad and random translation with norm 1 unit w.r.t. the wold frame; the second camera $\boldsymbol{C}'$ is obtained following a set of conditions which will be defined later and for each configuration we generate 10 random problem instances. We consider the pin-hole camera model and perturb the observations in each image by a Gaussian noise with zero mean and standard deviation $\sigma = 0.5, 1, \ldots, 8$. The image size is 1024^2 and the focal length is 512. We let the distance d_0 vary from $d_0 = 1, \ldots, 64$.

We evaluate our solver on five different configurations for the camera poses: (1) 'general': the second pose is generated as the first one; (2) 'lateral': the second camera has a relative translation w.r.t. the first one with form $\boldsymbol{t} = [a, b, 0]^T$ and zero rotation; (3) 'stereo': the translation from the first camera to the second has form $\boldsymbol{t} = [1, 0, 0]^T$ and zero rotation (simulating a rectified stereo setup); (4) 'diagonal': relative poses with translation $\boldsymbol{t} = 1/\sqrt{3}[1,1,1]^T$ and zero rotation; and (5) 'forward': the second camera has relative translation $\boldsymbol{t} = [0,0,1]^T$ and zero rotation. The cameras with fixed translation direction ('stereo', 'diagonal' and 'forward') are perturbed with uniform noise with standard deviation 0.01 and zero mean. This leads to more than one million of different triangulation problems.

Results for the primal-dual solver in algorithm (1): Figures (3) shows the errors between the observed points and the corrected ones for the noise levels $\sigma = 0.5, 2.5, 5.0, 8.0\,\mathtt{pix}$ (X-axis) and $d_0 = 4$ unit (first row) and $d_0 = 32$ units (second row). We consider the ℓ_1 (Figures (3f), (3f)), ℓ_2 (Figures (3g), (3g)) and ℓ_∞ (Figures (3h), (3h)) of the solution $\boldsymbol{w}$ for each solver (except SVD). Figures (3a), (3e) show the Euclidean distance between the reconstructed 3D points and the ground-truth data. Observe first that ITER, due to its iterative nature, get trapped into local minima hence returning poor solutions with large costs and errors. In general, the error in 2D are similar for all the solvers. The general HS-G attain lower errors in observation since the planar constraint is not imposed. However, its errors in 3D are larger than those for POLY and DUAL, and similar to the linear solver SVD. Notice that POLY and DUAL attain the same errors in all the graphics. Figure (6a) shows the error in the ℓ_2 sense between the corrected observations obtained with our method DUAL and the polynomial solver POLY as a cumulative distribution function. We include the results for all the experiments since the values of the difference were similar for all the considered noise level and distance d_0 to the plane. Observe that the graphic resembles a step, and for more than 80% of the problem instances the value of the error was below $5e-13$.

Computational time: We provide next the computational time required by each algorithm for all the synthetic experiments. The evaluation was performed in a desktop PC with CPU $i7-4702$MQ, 2.2GHz and 8 GB RAM. Additionally, we provide the distribution of times for all the configurations by all the algorithms except ITER in Figure 5. We don't include ITER since it is slower than the other options. The solvers required the same time for all the different camera configurations, and hence we provide the mean (avg) and standard deviation (std) for all of them. DUAL requires 3.1 microseconds (avg) with 2.2 microseconds as std, POLY goes to 7.7 microseconds (avg) and 3.4 microseconds (std). The iterative solver ITER, which usually takes 20 iterations, requires 1.3 milliseconds (avg) and 1.1 milliseconds (std). The general polynomial solver HS-G is similar to POLY, with (avg) 7.6 microseconds and (std) 3.9 microseconds. Our solver is then at least two times faster than POLY. On the other hand, solving the dual problem from scratch with SDPT3 as Interior Point Method (IPM) and CVX as modeling tool in matlab requires 0.27 seconds per problem instance, hence being slower than all the previous solvers. Note, though, that since our solver returns both primal and dual solutions, they can be leveraged as initial guess by these IPMs when our algorithm cannot certify optimality.

Now we break down the different stages of Alg. (1). We also observe here a similarity between the times for all the camera configurations and hence only provide the mean (avg) and standard deviation (std) for all of them. We consider the next five

Synthetic data: errors

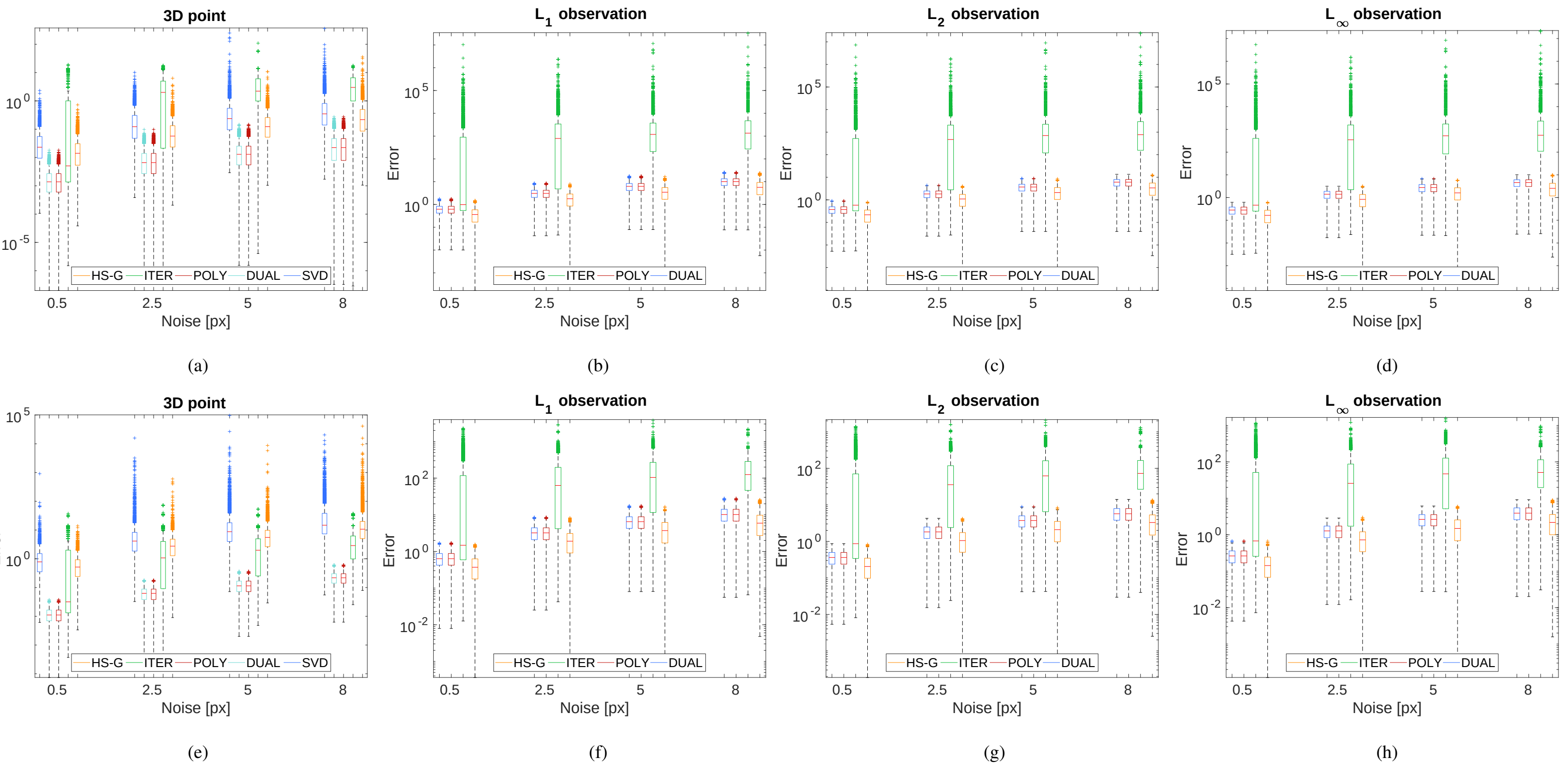


Fig. 3: **Synthetic data: errors** Errors (log-scale) of the corrected matches for the different considered solvers: first column shows the Euclidean 3D error for the reconstructed world point; and second, third and fourth columns show the 2D observation errors measured as ℓ_1, ℓ_2 and ℓ_∞, respectively for all the different solvers. First row has the experiments for a plane parallel to the x-y plane w.r.t. the world frame at distance to the origin $d_0 = 1$ unit and second row $d_0 = 32$ units.

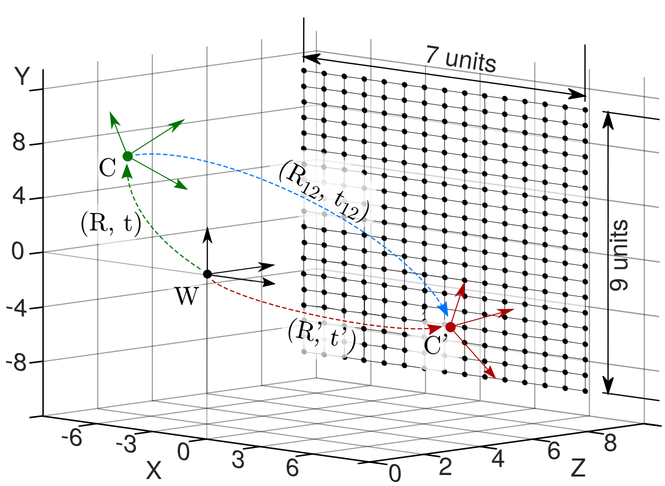


Fig. 4: Setup of the synthetic experiments with a plane of size 9×7 units and 285 points. The first camera's pose $\boldsymbol{C}$ has a random rotation $\boldsymbol{R}$ and translation $\boldsymbol{t}$ w.r.t. the world frame $\boldsymbol{W}$. The second camera's pose $\boldsymbol{C}'$ depends on the configuration of the experiment (more in the text).

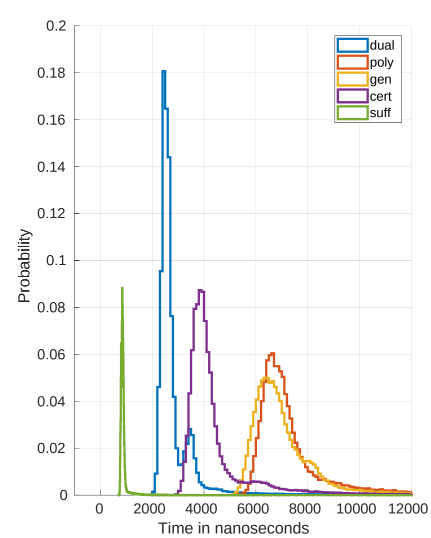
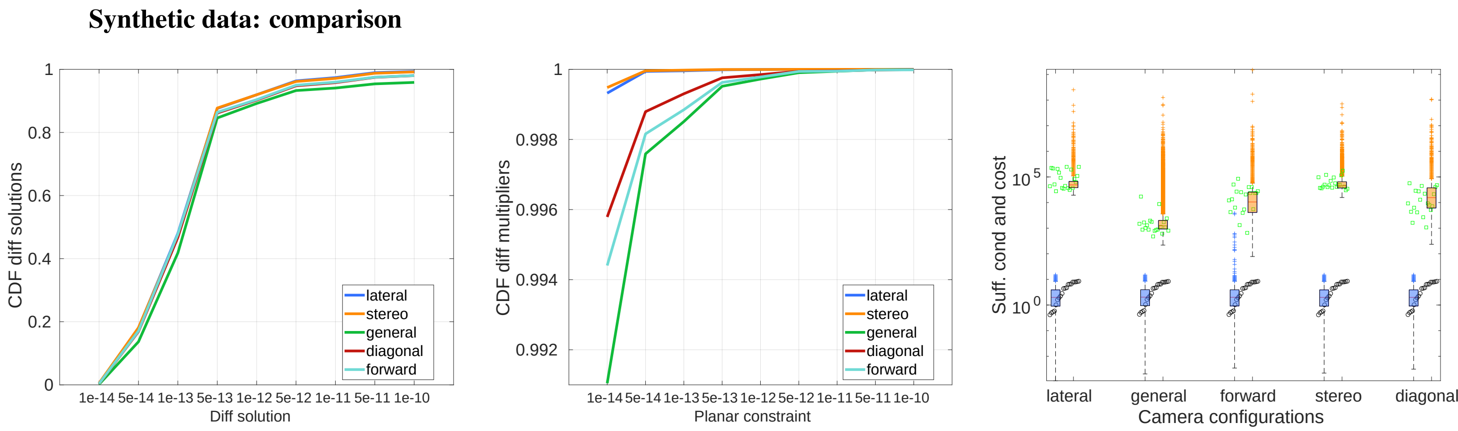


Fig. 5: Probability of the computational time (Y-axis) as a function of the nanoseconds required by the solvers (X-axis), including the certifier and sufficient condition (see legend). We don't include ITER for being slower than the other options. These results include all the configurations tested on the evaluation.



(a) (b) (c)

Fig. 6: **Synthetic data: comparison.** Figure (6a) depicts the cumulative distribution function (CDF) of the difference in the ℓ_2 sense between the solutions of POLY and DUAL. Figure (6b) shows the CDF for the multipliers obtained by Alg. (1) and Alg. (2) for the same problem instances. Figure (6c) depicts the cost of the optimal solution for DUAL (blue boxplot and black circles) and the upper bound estimated from Eq. (21) (orange boxplot and green squares).

stages of the Algorithm (all the values are in microseconds): (1) Problem initialization -line 1-: (avg) 0.3397, (std) 0.4307; (2) data initialization -lines $3-5$-: (avg) 0.4356, (std) 0.4456; (3) polynomial coefficients estimation -line 6-: (avg) 1.355, (std) 1.1899; (4) root finding -lines $7-11$-: (avg) 0.8671, (std) 0.6012; and (5) solution recovery -lines $12-18$-: (avg) 0.0574, (std) 0.2763. The root finding algorithm requires 3.5 iterations in average (wit std 2.4 iterations). Therefore, our algorithm can be improved in terms of computational time by implementing a faster root finding algorithm within the desired interval, which is not contemplated here.

Results for the certifiable algorithm (2): This section evaluates the accuracy of the certifiable algorithm in Alg. (2). Considering the previous results, we consider as ground-truth the primal-dual solution from Alg. (1). We perform the same evaluation with the solutions for POLY and obtain similar results. Figure (6b) shows the difference as CDF in the ℓ_2 sense between the multipliers estimated by Alg. (1) and Alg. (2) for the solution returned by the former. Observe that the difference remains under $5e-12$, and show that Alg. (2) is able to certify all the optimal solutions. Regarding computational time, the slowest stage in Alg. (2) is the estimation of the candidate to dual point $\hat{\lambda}$ with Eq. (20). In total, the certification requires 2.3 microseconds (avg) and 1 microsecond (std). Recall that this Algorithm *only* certifies a solution, it does not obtain it.

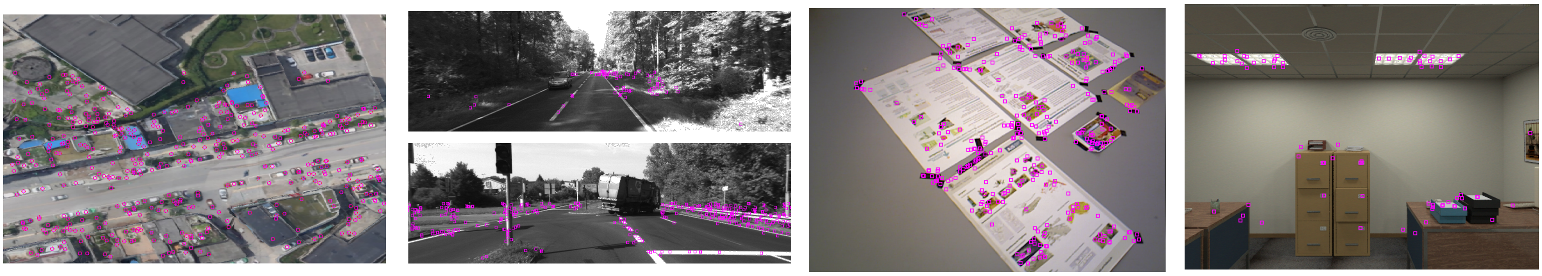
Results for the sufficient condition: To conclude this part of the evaluation, we test the performance of the sufficient condition in Eq. (21). The condition was applied to all the solutions returned by DUAL, and 100% of the cases the condition was able to estimate optimality, *i.e.*, the sufficient condition is tight in practice. Figure (6c) shows the bound on the cost derived from the sufficient condition (orange boxplot) and the cost attained by the global optimum of the problem (blue boxplot). We also sample randomly some of the results and plot the specific values as circles (green for the condition and black for the cost). Notice that the sufficient condition is tight in practice. Checking the sufficient condition requires less than 1 microsecond per problem instance.

6.2. *Evaluation on real data*

We conclude the evaluation of our algorithms with an extensive set of experiments on real data. We select the next datasets that provide with different camera configurations and scenes that are found in real-world applications:

1. Ground images from an aerial vehicle [dataset] Liu and Ji (2020). The images have texture with many buildings and roads. The buildings' height are negligible w.r.t. the ground plane and the whole scene can be approximated by a single plane with known equation. We use the next sequences, each of them with 200 pair of images: VAL-009-77, denoted by AERIAL: 009-77; VAL-013-68, denoted by AERIAL: 013-08; TEST-002-50, denoted by AERIAL: 002-50; TEST-003-73, denoted by AERIAL: 003-73; TEST-009-67, denoted by AERIAL: 009-67; and TEST-011-38, denoted by AERIAL: 011-38.
2. Road images from a car [dataset] Andreas Geiger et al. (2013). The sequences were recorded by a moving vehicle with a forward pointing camera. The road is visible in all the images and presents some strong features. The next sequences are considered: from the KITTI dataset: ROAD-2011-09-26-DRIVE-0027-SYNC (185 pair of images), denoted by KITTI: ROAD-27; ROAD-2011-09-30-DRIVE-0016-SYNC (277 pair of images), denoted by KITTI: ROAD-30; ROAD-2011-09-26-DRIVE-0029-SYNC (428 pair of images), denoted by KITTI: ROAD-29; and ROAD-2011-09-26-DRIVE-0032-SYNC (390 pair of images), denoted by KITTI: ROAD-32.
3. Indoor scenes [dataset] J. Sturm et al. (2012). The images show a set of planes (at least one in all the sequences) with some texture provided by a set of posters. The ground plane, when visible, does not provide any suitable feature and we rely only on the posters for the triangulation. We consider the next sequences: FREIBURG3-NOSTRUCTURE-TEXTURE-FAR (73 pair of images)

Real data: samples



(a) (b) (c) (d)

Fig. 7: **Real data: samples** Figures (7a)-(7d) show sample images from the considered datasets with the set of correspondences marked in magenta circles: Aerial: 009-77 in Fig. (7a); Fig. (7b) -top- KITTI: road-27 and -bottom- KITTI: road-29; TUM-struct-far in Fig. (7c); and of kt1 in Fig. (7d).

denoted by TUM-nonstruct-far; freiburg3-nostructure-texture-near-withloop (326 pair of images) denoted by TUM-nonstruct-near; freiburg3-structure-texture-far (180 pair of images) denoted by TUM-struct-far; and freiburg3-structure-texture-near (210 pair of images) denoted by TUM-struct-near.

4. Interior building images [dataset] A. Handa et al. (2014). The sequences show two different indoor scenes with texture and structure. Strong features on planar surfaces appear several times on the images. Floor and walls also provide some features. We consider the next four sequences: lr kt1 (321 images) (denoted by the same name); lr kt2 (292 images); of kt1 (321 images); and of kt3 (413 images).

Samples of the sequences with the set of correspondences marked in magenta are shown in shown in Figure (7). All datasets provide us with ground-truth poses, point clouds (or depth images) and intrinsic calibration data, and a total of 4615 pair of images. To generate correspondences we extract and match SURF features Bay et al. (2008), filtering those that do not lie on the dominant planes. This also allows us to discard outliers. We correct the matches with the above-mentioned solvers.

Figures (8e)-(8h) depict the ℓ_2 norm of the correction for the different solvers and datasets following the above-mentioned groups. Due to space limits and the similarity of the results, we move the errors in the ℓ_1 and ℓ_∞ sense to the *Supplementary material* Section (Appendix C). Notice that DUAL and POLY attain the same 2D error, while ITER is unstable for some problem instances, as it was observed in the synthetic evaluation. Notice also that the general solver HS-G also becomes unstable, which appears as outliers in the graphics. We test the similarity between the solutions from DUAL and POLY by measuring the difference in the ℓ_2 sense. We obtain a similar distribution of the errors than in the synthetic evaluation. The graphics can be found in the *Supplementary material* Section (Appendix C).

We also measure the difference between the lagrange multipliers obtained by Alg. (1) and the certifier in Alg. (2) and observe that the multipliers agree with at least $1e-15$ of accuracy in all the cases, and so the certifier is able to certify all the solutions to the non-convex problem (O). Last, we include here the results for the bound in Figures (8e)- (8h) derive from the sufficient condition in Eq. (21) following the same format than in Figure (6c). Notice that for these problem instances the sufficient condition is also able to certify all the optimal solution as such. For the aerial sequences in Figure (8e) the upper bound is $\sigma(\boldsymbol{B})/0$.

7. Conclusions and future work

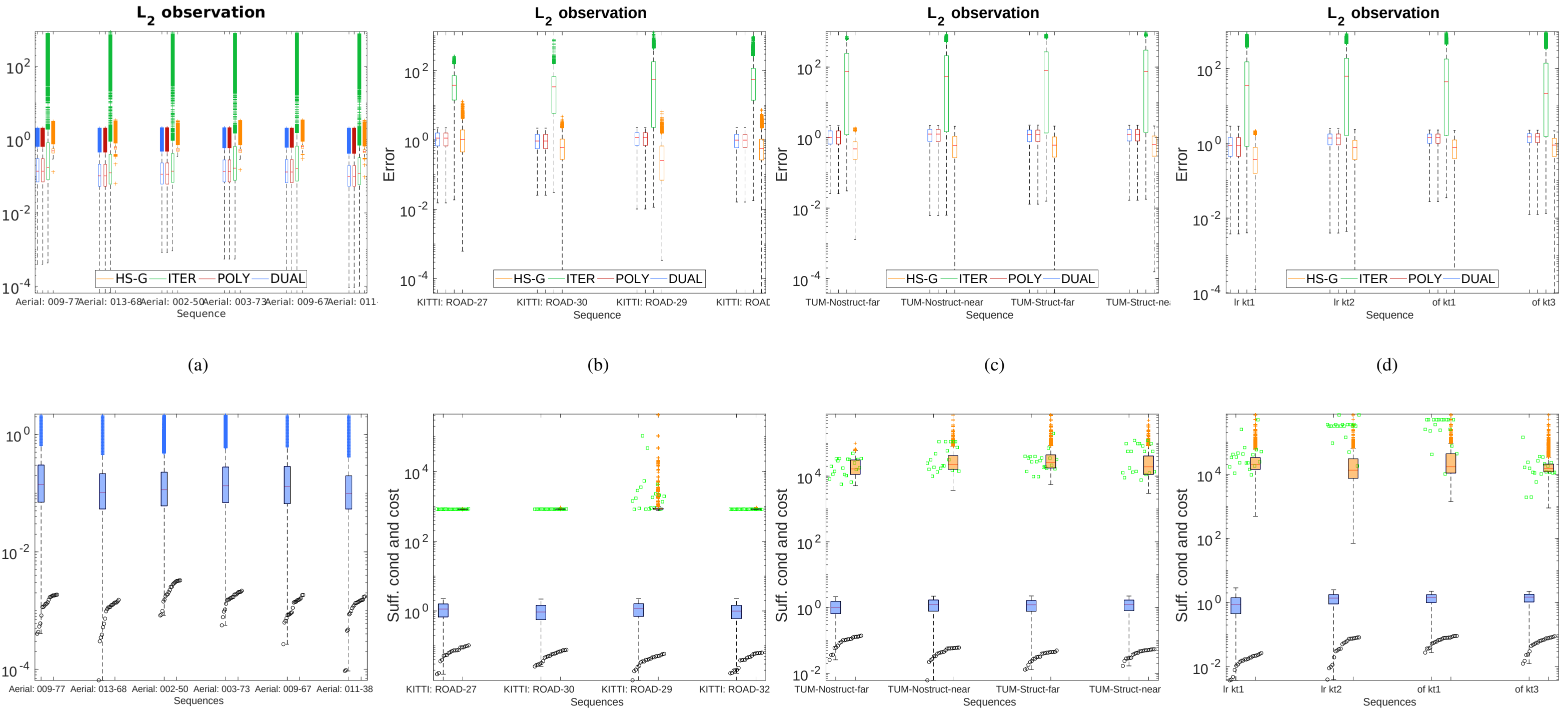
In this work, we have tackled the two-view triangulation problem under the restriction that the 3D unknown world points must lie on a plane. Based on duality theory, we have proposed a primal-dual solver that is able to both estimate and certify the solution to this non-convex problem. The solver is insensible to high noise and always return a feasible solution. It is also faster than the other certifiable algorithm in Chum et al. (2005), and the difference between the solutions remains under $1e-11$ for most problem instances. From the formulation of this solver, we have derived an optimality certificate that allows to certify a given feasible solution for the original problem. This certifier is shown to work in practice and is able to certify all the solutions in 2.3 microseconds. Moreover, based on this certifier we derived a sufficient (but not necessary) condition for optimality that also works in practice, taking less than one microsecond to certify solutions. We carried out extensive experiments on both synthetic and real data that prove the utility of our proposals. The code is made available at `https://github.com/mergarsal`.

As future work, we devised to extend this approach to more than two views and to include these certifiable solvers into complete pipelines. We also contemplate to prove theoretically than the proposed relaxation is tight even with highly noise observations.

Acknowledgments

This work has been supported by the grant program FPU18/01526 and the research projects WISER (DPI2017-84827-R) and ARPEGGIO (PID2020-117057), all funded by the Spanish Government. We also want to thank two anonymous reviewers for their comments and suggestions that help us improve the quality of this manuscript.

Real data: results



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 8: **Real data: results** Figures (8a)-(8d) show the ℓ_2 norm (log-scale) of the correction for the different solvers and datasets (see legends and text). Figures (8e)-(8h) depict the cost of the optimal solution for DUAL (blue boxplot and black circles) and the upper bound estimated from Eq. (21) (orange boxplot and green squares). For the aerial images (Fig. (8e)), the upper bound has the form $\sigma(\boldsymbol{B})/0$.

References

[dataset] A. Handa, Whelan, T., McDonald, J., Davison, A., 2014. A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM, in: IEEE Intl. Conf. on Robotics and Automation, ICRA, Hong Kong, China.

Aholt, C., Agarwal, S., Thomas, R., 2012. A qcqp approach to triangulation, in: European Conference on Computer Vision, Springer. pp. 654–667.

[dataset] Andreas Geiger, Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets robotics: The kitti dataset. International Journal of Robotics Research (IJRR) .

Arndt, C., Sabzevari, R., Civera, J., . From points to planes-adding planar constraints to monocular slam factor graphs .

Artieda, J., Sebastian, J.M., Campoy, P., Correa, J.F., Mondragón, I.F., Martínez, C., Olivares, M., 2009. Visual 3-d slam from uavs. Journal of Intelligent and Robotic Systems 55, 299–321.

Bartoli, A., Sturm, P., 2001. Constrained structure and motion from n views of a piecewise planar scene, in: Virtual and Augmented Architecture (VAA'01). Springer, pp. 195–206.

Bartoli, A., Sturm, P., 2003. Constrained structure and motion from multiple uncalibrated views of a piecewise planar scene. International Journal of Computer Vision 52, 45–64.

Bartoli, A., Sturm, P., Haraud, R., 2001. Projective structure and motion from two views of a piecewise planar scene, in: Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, IEEE. pp. 593–598.

Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-up robust features (surf). Computer vision and image understanding 110, 346–359.

Beardsley, P.A., Zisserman, A., Murray, D.W., 1994. Navigation using affine structure from motion, in: European Conference on Computer Vision, Springer. pp. 85–96.

Boyd, S., Vandenberghe, L., 2004. Convex optimization. Cambridge university press.

Briales, J., Gonzalez-Jimenez, J., 2016. Fast global optimality verification in 3d slam, in: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE. pp. 4630–4636.

Briales, J., Gonzalez-Jimenez, J., 2017a. Cartan-sync: Fast and global se (d)-synchronization. IEEE Robotics and Automation Letters 2, 2127–2134.

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

Caballero, F., Merino, L., Ferruz, J., Ollero, A., 2009. Vision-based odometry and slam for medium and high altitude flying uavs. Journal of Intelligent and Robotic Systems 54, 137–161.

Carlone, L., Dellaert, F., 2015. Duality-based verification techniques for 2d slam, in: 2015 IEEE international conference on robotics and automation (ICRA), IEEE. pp. 4589–4596.

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

Chum, O., Pajdla, T., Sturm, P., 2005. The geometric error for homographies. Computer Vision and Image Understanding 97, 86–102.

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

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

Faugeras, O.D., Lustman, F., 1988. Motion and structure from motion in a piecewise planar environment. International Journal of Pattern Recognition and Artificial Intelligence 2, 485–508.

Garcia-Salguero, M., Briales, J., Gonzalez-Jimenez, J., 2021. Certifiable relative pose estimation. Image and Vision Computing 109, 104142.

Gomez-Ojeda, R., Moreno, F.A., Zuniga-Noël, D., Scaramuzza, D., Gonzalez-Jimenez, J., 2019. Pl-slam: A stereo slam system through the combination of points and line segments. IEEE Transactions on Robotics 35, 734–746.

Hartley, R., Seo, Y., 2008. Verifying global minima for l 2 minimization problems, in: 2008 IEEE Conference on Computer Vision and Pattern Recognition, IEEE. pp. 1–8.

Hartley, R., Zisserman, A., 2003. Multiple view geometry in computer vision.

Cambridge university press.
Hartley, R.I., Sturm, P., 1997. Triangulation. Computer vision and image understanding 68, 146–157.
Hmam, H., 2010. Quadratic optimisation with one quadratic equality constraint. Technical Report. DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION EDINBURGH (AUSTRALIA) ELECTRONIC . . . .
Hoegner, L., Tuttas, S., Stilla, U., 2016. 3d building reconstruction and construction site monitoring from rgb and tir image sets, in: 2016 12th IEEE International Symposium on Electronics and Telecommunications (ISETC), IEEE. pp. 305–308.
Iglesias, J.P., Olsson, C., Kahl, F., 2020. Global optimality for point set registration using semidefinite programming, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8287–8295.
[dataset] J. Sturm, Engelhard, N., Endres, F., Burgard, W., Cremers, D., 2012. A benchmark for the evaluation of rgb-d slam systems, in: Proc. of the International Conference on Intelligent Robot Systems (IROS).
Jurevičius, R., Marcinkevičius, V., 2019. A data set of aerial imagery from robotics simulator for map-based localization systems benchmark. International Journal of Intelligent Unmanned Systems .
Kanatani, K., Niitsuma, H., 2011. Optimal two-view planar scene triangulation. IPSJ Transactions on Computer Vision and Applications 3, 67–79.
Kanatani, K., Sugaya, Y., Niitsuma, H., 2008. Triangulation from two views revisited: Hartley-sturm vs. optimal correction. practice 4.
Kanazawa, Y., Kanatani, K., 1995. Direct reconstruction of planar surfaces by stereo vision. IEICE TRANSACTIONS on Information and Systems 78, 917–922.
Kang, Z., Yang, J., Yang, Z., Cheng, S., 2020. A review of techniques for 3d reconstruction of indoor environments. ISPRS International Journal of Geo-Information 9, 330.
Karrer, M., Chli, M., 2020. Distributed variable-baseline stereo slam from two uavs. arXiv preprint arXiv:2009.04801 .
Koch, D., Despotovic, M., Leiber, S., Sakeena, M., Döller, M., Zeppelzauer, M., 2019. Real estate image analysis: A literature review. Journal of Real Estate Literature 27, 269–300.
Lasserre, J.B., 2008. A semidefinite programming approach to the generalized problem of moments. Mathematical Programming 112, 65–92.
Laveau, O.F.S., Robert, L., Czurka, G., Zeller, C., 1998. 3d reconstruction of urban scenes from sequences of images. Computer Vision and Image Understanding 69, 292–309.
Lee, S.H., Civera, J., 2019. Triangulation: Why optimize? British Machine Vision Conference (BMVC) .
Lindstrom, P., 2010. Triangulation made easy, in: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, IEEE. pp. 1554–1561.
[dataset] Liu, J., Ji, S., 2020. A novel recurrent encoder-decoder structure for large-scale multi-view stereo reconstruction from an open aerial dataset, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6050–6059.
Lu, Y., Xue, Z., Xia, G.S., Zhang, L., 2018. A survey on vision-based uav navigation. Geo-spatial information science 21, 21–32.
Micusik, B., Wildenauer, H., 2017. Plane refined structure from motion, in: Scandinavian Conference on Image Analysis, Springer. pp. 29–40.
Mur-Artal, R., Montiel, J.M.M., Tardos, J.D., 2015. Orb-slam: a versatile and accurate monocular slam system. IEEE transactions on robotics 31, 1147–1163.
Poullis, C., You, S., 2011. 3d reconstruction of urban areas, in: 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, IEEE. pp. 33–40.
Rosen, D.M., Carlone, L., Bandeira, A.S., Leonard, J.J., 2019. Se-sync: A certifiably correct algorithm for synchronization over the special euclidean group. The International Journal of Robotics Research 38, 95–125.
Sportouche, H., Tupin, F., Denise, L., 2009. Building extraction and 3d reconstruction in urban areas from high-resolution optical and sar imagery, in: 2009 Joint Urban Remote Sensing Event, IEEE. pp. 1–11.
Sturm, J.F., 1999. Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones. Optimization methods and software 11, 625–653.
Szeliski, R., Torr, P.H., 1998. Geometrically constrained structure from motion: Points on planes, in: European Workshop on 3D Structure from Multiple Images of Large-Scale Environments, Springer. pp. 171–186.
Toh, K.C., Todd, M.J., Tütüncü, R.H., 1999. Sdpt3—a matlab software package for semidefinite programming, version 1.3. Optimization methods and software 11, 545–581.
Triggs, B., McLauchlan, P.F., Hartley, R.I., Fitzgibbon, A.W., 1999. Bundle adjustment—a modern synthesis, in: International workshop on vision algorithms, Springer. pp. 298–372.
Wefelscheid, C., Hänsch, R., Hellwich, O., 2011. Three-dimensional building reconstruction using images obtained by unmanned aerial vehicles. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38.
Yang, H., Carlone, L., 2019. A quaternion-based certifiably optimal solution to the wahba problem with outliers, in: Proceedings of the IEEE International Conference on Computer Vision, pp. 1665–1674.
Yang, H., Shi, J., Carlone, L., 2020. Teaser: Fast and certifiable point cloud registration. arXiv preprint arXiv:2001.07715 .
Yang, K., Fang, W., Zhao, Y., Deng, N., 2019. Iteratively reweighted midpoint method for fast multiple view triangulation. IEEE Robotics and Automation Letters 4, 708–715.
Zhao, J., 2020. An efficient solution to non-minimal case essential matrix estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence doi:10.1109/TPAMI.2020.3030161.
Zhao, J., Xu, W., Kneip, L., 2020. A certifiably globally optimal solution to generalized essential matrix estimation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12034–12043.

Appendix A. Dual problem for the non-convex problem (QCQP)

In this Section we provide the derivation of the set of conditions in Eqs. (5)-(9). For that, we follow the standard procedure for QCQPs with the form in Prob. QCQP that can be found in the literature and textbooks, *e.g.* Boyd and Vandenberghe (2004). We then restrict ourselves to problems with the specific form and refer the reader to this reference for further information.

The dual problem for the primal QCQP is defined as:

$$g^\star = \max_{\lambda_1,\lambda_2,\rho\in\mathbb{R}} \min_{\boldsymbol{x}\in\mathbb{R}^{2N}} \mathcal{L}(\boldsymbol{x},\lambda_1,\lambda_2,\rho), \tag{A.1}$$

where $\mathcal{L}(\boldsymbol{x},\lambda_1,\lambda_2,\rho)$ is the Lagrangian

$$\mathcal{L}(\boldsymbol{x},\lambda_1,\lambda_2,\rho) = \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{x}^T\lambda_1\boldsymbol{B}_1\boldsymbol{x} - \boldsymbol{x}^T\lambda_2\boldsymbol{B}_2\boldsymbol{x} - \rho\boldsymbol{x}^T\boldsymbol{L}\boldsymbol{x} + \rho = \tag{A.2}$$

$$= \boldsymbol{x}^T(\boldsymbol{Q} - \lambda_1\boldsymbol{B}_1 - \lambda_2\boldsymbol{B}_2 - \rho\boldsymbol{L})\boldsymbol{x} + \rho. \tag{A.3}$$

The dual function $d(\lambda_1,\lambda_2,\rho) = \min_{\boldsymbol{x}\in\mathbb{R}^{2N}}\mathcal{L}(\boldsymbol{x},\lambda_1,\lambda_2,\rho)$ has finite minimum *iff* the Hessian of the Lagrangian $\boldsymbol{H}_\lambda \doteq \boldsymbol{Q} - \lambda_1\boldsymbol{B}_1 - \lambda_2\boldsymbol{B}_2 - \rho\boldsymbol{L}$ is semidefinite positive, *i.e.* $\boldsymbol{H}_\lambda \succeq 0$. In this case, the minimum value of the expression $\boldsymbol{x}^T\boldsymbol{H}_\lambda\boldsymbol{x}$ is zero, and the Lagrangian takes ρ as minimum cost. Since we are interested in finite values for the dual problem, we can restrict our development to this case, and so the dual problem has the final form

$$g^\star = \max_{\lambda_1,\lambda_2,\rho\in\mathbb{R}} \rho \quad \text{subject to } \boldsymbol{H}_\lambda \doteq \boldsymbol{Q} - \lambda_1\boldsymbol{B}_1 - \lambda_2\boldsymbol{B}_2 - \rho\boldsymbol{L} \succeq 0 \tag{D}$$

where the symbol $\succ$ (resp. $\succeq$) implies positive definite PD (resp. positive semidefinite PSD). Considering the form of the matrices $\boldsymbol{B}_1, \boldsymbol{B}_2$ in Eq. (4) and that $\boldsymbol{L}$ is a zero matrix with the bottom, right corner entry to one, we have that the Hessian of the Lagrangian $\boldsymbol{H}_\lambda$ has the explicit form

$$\boldsymbol{H}_\lambda = \begin{pmatrix} \boldsymbol{I}_4 - \lambda_1\boldsymbol{A}_1 - \lambda_2\boldsymbol{A}_2 & -(\lambda_1\boldsymbol{b}_1 + \lambda_2\boldsymbol{b}_2) \\ -(\lambda_1\boldsymbol{b}_1 + \lambda_2\boldsymbol{b}_2)^T & \rho - \lambda_1 c_1 - \lambda_2 c_2 \end{pmatrix}, \tag{A.4}$$

and by the Schur's complement of the block matrix $\boldsymbol{H}_\lambda$ (Boyd and Vandenberghe, 2004, Sec. A.5.5) we can re-write the condition that $\boldsymbol{H}_\lambda \succeq 0$ as the next three relations

$$\boldsymbol{S}_\lambda \doteq \boldsymbol{I}_4 - \lambda_1\boldsymbol{A}_1 - \lambda_2\boldsymbol{A}_2 \succeq 0 \tag{A.5}$$

$$-(\lambda_1\boldsymbol{b}_1 + \lambda_2\boldsymbol{b}_2)^T\boldsymbol{S}_\lambda^{-1}(\lambda_1\boldsymbol{b}_1 + \lambda_2\boldsymbol{b}_2) - \lambda_1 c_1 - \lambda_2 c_2 \geq \rho \tag{A.6}$$

$$\lambda_1\boldsymbol{b}_1 + \lambda_2\boldsymbol{b}_2 \in \mathcal{R}(\boldsymbol{I}_4 - \lambda_1\boldsymbol{A}_1 - \lambda_2\boldsymbol{A}_2), \tag{A.7}$$

being $\mathcal{R}(\boldsymbol{A})$ the range of the matrix $\boldsymbol{A}$.

Strong duality: Since the dual problem (D) is a relaxation of the primal (QCQP), the chain of inequalities holds

$$g(\lambda_1,\lambda_2,\rho) = \rho \leq g^\star \leq f^\star \leq f(\boldsymbol{w}) \tag{A.8}$$

for any primal $\boldsymbol{w}$ and dual $(\lambda_1,\lambda_2,\rho)$ feasible points, *i.e.* points in the domain of the problems (QCQP) and (D), respectively.

If the relaxation is *tight*, then $g^\star = f^\star$ and the dual (convex) problem permits us to certify the solution of the primal. We are thus interested in finding a pair of primal-dual feasible points $(\boldsymbol{x},\lambda_1,\lambda_2,\rho)$ such that the costs attained by them are equal. These solutions will be the global optimum of the respective problems, *and* provide us with an optimality certificate. In this case we have that

$$\boldsymbol{x}^{\star T}\boldsymbol{H}_\lambda^\star\boldsymbol{x}^\star = 0 \Leftrightarrow \boldsymbol{H}_\lambda^\star\boldsymbol{x}^\star = \boldsymbol{0}_{5\times1}, \tag{A.9}$$

where $\boldsymbol{H}_\lambda^\star$ is the Hessian matrix $\boldsymbol{H}_\lambda$ evaluated at the optimal (and feasible) points $(\lambda_1^\star,\lambda_2^\star,\rho^\star)$. Since $\boldsymbol{x}^\star$ is feasible for the primal, we know that its last entry is one WLOG[3] and we can re-write condition (A.9) as the two relations:

$$\boldsymbol{S}_\lambda^\star\boldsymbol{w}^\star = (\boldsymbol{I}_4 - \lambda_1^\star\boldsymbol{A}_1 - \lambda_2^\star\boldsymbol{A}_2)\boldsymbol{w}^\star = (\lambda_1^\star\boldsymbol{b}_1 + \lambda_2^\star\boldsymbol{b}_2) \tag{A.10}$$

$$\rho = -\lambda_1^\star c_1 - \lambda_2^\star c_2 - (\lambda_1^\star\boldsymbol{b}_1 + \lambda_2^\star\boldsymbol{b}_2)^T\boldsymbol{w}^\star = \boldsymbol{w}^{\star T}\boldsymbol{w}^\star \tag{A.11}$$

where the last equality in (A.11) follows from $\boldsymbol{w}^\star$ being primal feasible. Collecting all these conditions leads up to the set in Eqs. (5)-(9)

Appendix B. Shor's relaxation for the non-convex problem (QCQP)

An alternative convex relaxation for the primal (QCQP) is given by Shor's relaxation

$$h^\star = \min_{\boldsymbol{X}\in\mathbb{R}^5} \operatorname{tr}(\boldsymbol{Q}\boldsymbol{X}) \quad \text{subject to } \begin{array}{l} \operatorname{tr}(\boldsymbol{B}_1\boldsymbol{X}) = 0 \\ \operatorname{tr}(\boldsymbol{B}_2\boldsymbol{X}) = 0 \\ \operatorname{tr}(\boldsymbol{L}\boldsymbol{X}) = 1 \end{array} \tag{SDP}$$

Problem (SDP) is an instance of a semidefinite problem (SDP) which is convex by construction and can be solved by off-the-shelf tools in polynomial time, *e.g.* SeDuMi Sturm (1999) and SDPT3 Toh et al. (1999). Provided that some spectral conditions on the solution $\boldsymbol{X}^\star$ hold, we can recover and certify the optimal solution of the non-convex (QCQP) from $\boldsymbol{X}^\star$. In this work, we leverage instead the dual (D), which both estimates and certifies the optimum for (QCQP) in a simpler and more efficient way.

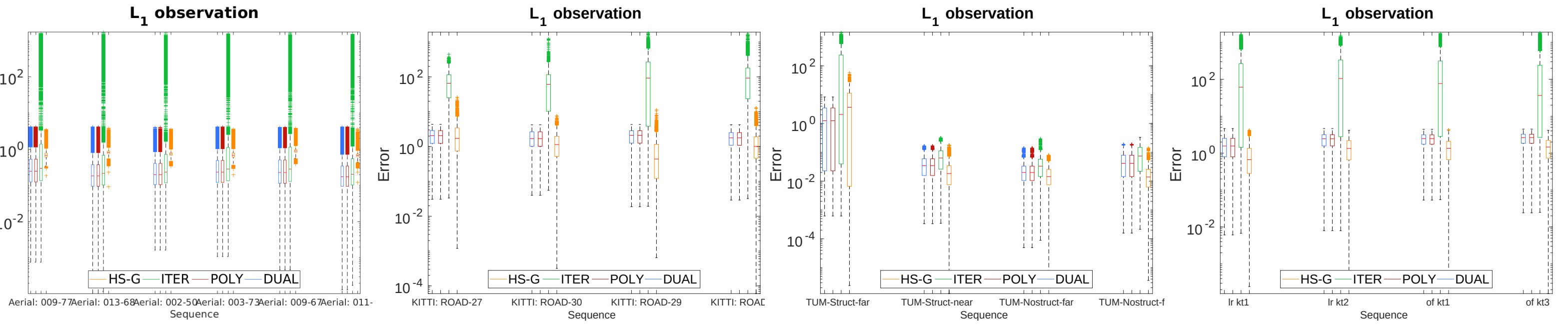
Appendix C. Further results of real data

In this Section we include the graphics for the evaluation on real data that due to space limits and/or similarity with other results were not included in the manuscript.

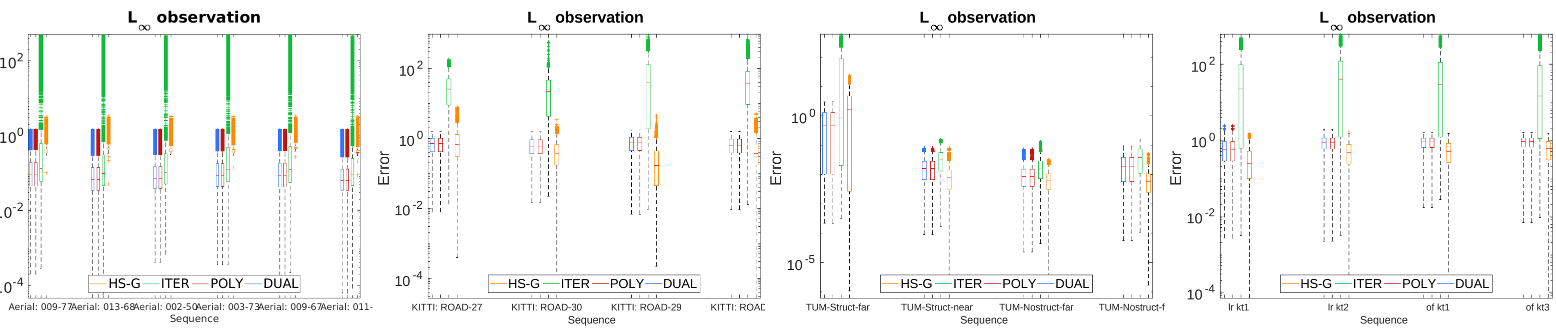
Figures (C.9a)-(C.9d) show the ℓ_1 norm (log-scale) of the correction for the different solvers and datasets. Figures (C.9e)-(C.9h) depict the same results under the ℓ_∞ norm. Notice that ITER and HS-G become unstable for some problem instances, while DUAL and POLY attain always the same errors.

[3] Since $+\boldsymbol{x}$ and $-\boldsymbol{x}$ are solutions to the (QCQP), we can select the one with the last entry equal to +1

Real data: further results



(a) (b) (c) (d)



(e) (f) (g) (h)

Fig. C.9: **Real data: further results**: Figures (C.9a)-(C.9d) show the ℓ_1 norm (log-scale) of the correction for the different solvers and datasets. Figures (C.9e)-(C.9h) show the results for the ℓ_∞ norm.

Figures (C.10a)- (C.10d) show the difference as CDF between the solutions of POLY and DUAL for the different datasets and sequences. These errors are similar to the ones on synthetic data in Figure (6a) and are all below $5e-11$. We also measure the value of the homography constraint in the ℓ_2 sense for the solution returned by DUAL. Figures (C.10e)- (C.10h) depict the cumulative distribution function of this value for the different sequences. Observe that for all the problem instances the value of constraint remains under $1e-10$.

Real data: comparison

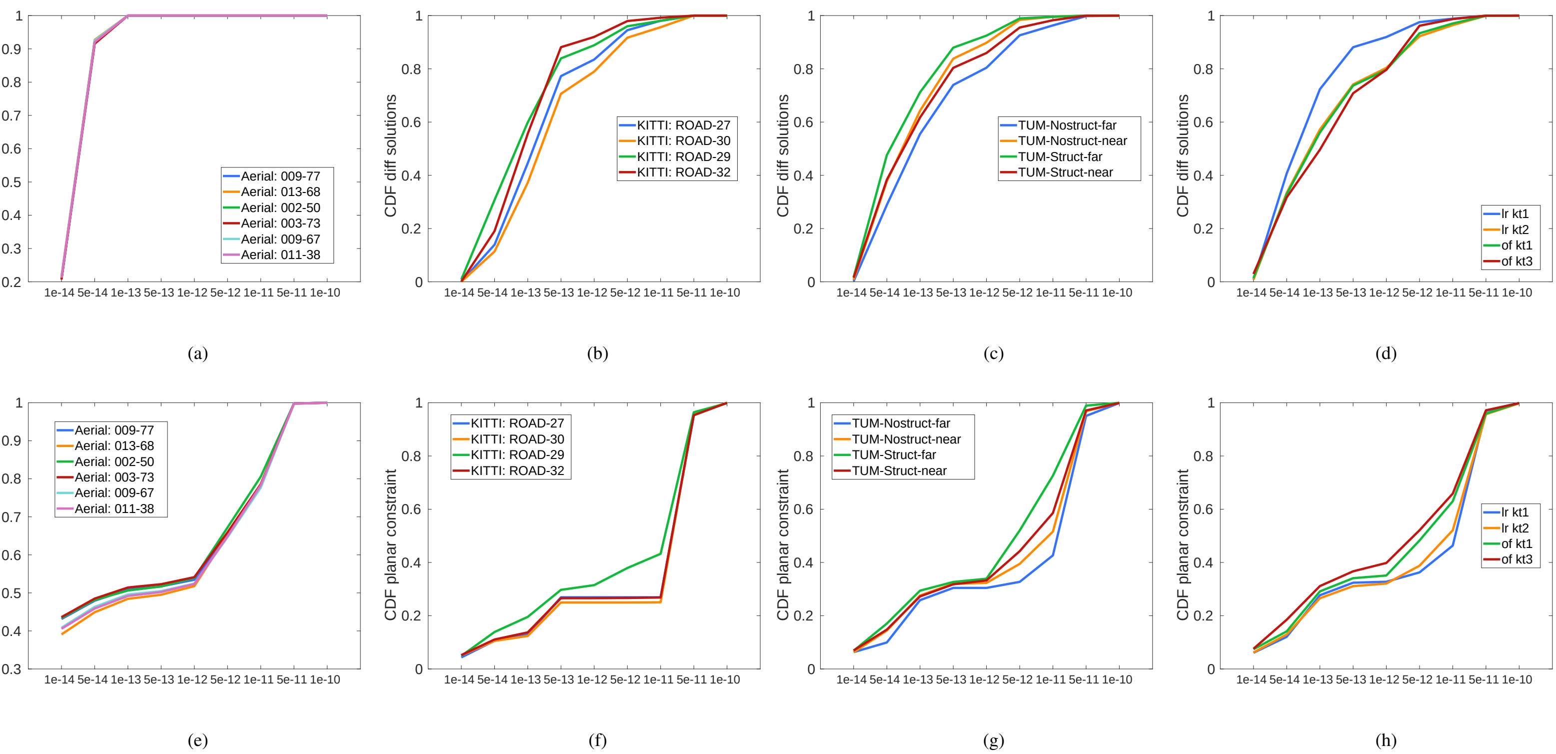


Fig. C.10: **Real data: comparison** First row shows the cumulative distribution function (CDF) of the difference between the solutions returned by DUAL and POLY for the different datasets and sequences (see X-axis). The second row shows the CDF of the homography constraint (in the ℓ_2 sense) for the solution returned by DUAL.