A QCQP Approach to Triangulation

Chris Aholt[1], Sameer Agarwal[2], and Rekha Thomas[1]

[1] University of Washington
[2] Google Inc.

Abstract. Triangulation of a three-dimensional point from $n \geq 2$ two-dimensional images can be formulated as a quadratically constrained quadratic program. We propose an algorithm to extract candidate solutions to this problem from its semidefinite programming relaxations. We then describe a sufficient condition and a polynomial time test for certifying when such a solution is optimal. This test has no false positives. Experiments indicate that false negatives are rare, and the algorithm has excellent performance in practice. We explain this phenomenon in terms of the geometry of the triangulation problem.

1 Introduction

We consider the problem of triangulating a point $X \in \mathbb{R}^3$ from $n \geq 2$ noisy image projections. This is a fundamental problem in multi-view geometry and is a crucial subroutine in all structure-from-motion systems [4].

Formally, let the point $X \in \mathbb{R}^3$ be visible in $n \geq 2$ images. Also let $P_i \in \mathbb{R}^{3\times 4}$ be a projective camera and $x_i \in \mathbb{R}^2$ be the projection of X in image i, i.e,

$$x_i = \Pi P_i \tilde{X}, \ \forall \ i = 1, \ldots, n, \tag{1}$$

where, using `MATLAB` notation, $\tilde{X} = \left[X; 1\right]$ and $\Pi \left[u; v; w\right] = \left[u/w; v/w\right]$.

Given the set $\{x_i\}$ of noise free projections, it is easy to determine X using a linear algorithm based on singular value decomposition (SVD) [4]. However, in practice we are given $\widehat{x}_i = x_i + \eta_i$, where η_i is noise, and there are no guarantees on the quality of the solution returned by the linear algorithm.

For simplicity, we assume that $\eta_i \sim \mathcal{N}(0, \sigma I)$. Then the *triangulation problem* is to find the maximum likelihood estimate of X given the noisy observations $\{\widehat{x}_i\}$. Assuming such a point X always exists, this is equivalent to solving:

$$\arg\min_X \sum_i^n \|\Pi P_i \tilde{X} - \widehat{x}_i\|^2. \tag{2}$$

Here and in the rest of the paper we will ignore the constraint that the point X has positive depth in each image. The above optimization problem is an instance of fractional programming which is in general hard [1]. An efficient and optimal solution of (2) is the subject of this paper.

For $n = 2$, Hartley & Sturm showed that (2) can be solved optimally in polynomial time [3]. For $n = 3$, a Gröbner basis based algorithm for (2) was proposed in [20]. This algorithm relies on the observation that generically, the optimal solution to (2) is one among the 47 solutions to a certain system of polynomial equations. This Gröbner basis method is not usefully extendable to higher n and efficient optimal triangulation for $n \geq 4$ views has remained an unsolved problem. Other approaches either use the linear SVD based algorithm as initialization followed by non-linear refinement which lead to locally optimal solutions with no guarantees on the run time complexity [4], or optimal algorithms whose worst case complexity is exponential in n [7, 8, 13].

We present a new triangulation algorithm for $n \geq 2$ views. Based on semidefinite programming, the algorithm in polynomial time either determines a globally optimal solution to (2) or informs the user that it is unable to do so. Theoretically, the operating range (in terms of image noise) of the algorithm is limited and depends on the particular configuration of the cameras. In practice our method computes the optimal solution in the vast majority of test cases. In the rare case that optimality cannot be certified, the algorithm returns a solution which can be used as an initializer for nonlinear least squares iteration.

The paper is organized as follows. In Section 2 we formulate triangulation as a constrained quadratic optimization problem. We present semidefinite relaxations to this problem in Section 3. We propose our triangulation algorithm in Section 4 and analyze its performance. We also provide theoretical explanation for why the algorithm works. Section 5 presents experiments on synthetic and real data and we conclude in Section 6 with a discussion.

Notation. We will use `MATLAB` notation to manipulate matrices and vectors, e.g., $A[1:2, 2:3]$ refers to a 2×2 submatrix of A. $P = \{P_1, \ldots, P_n\}$ denotes the set of cameras. $x = [x_1; \ldots; x_n]$ denotes a vector of image points, one in each camera, and $\widehat{x} = [\widehat{x}_1; \ldots; \widehat{x}_n]$ denotes the vector of image observations. Both x and $\widehat{x}$ lie in $\mathbb{R}^{2n}$. If $y \in \mathbb{R}^m$ is a vector, then $\tilde{y} = \big[y; 1\big]$ is the homogenized version of y. The inner product space of $k \times k$ real symmetric matrices is denoted $\mathcal{S}^k$ with the inner product $\langle A, B \rangle = \sum_{1 \leq i,j \leq k} A_{ij} B_{ij}$. The set $\mathcal{S}^k_+ \subseteq \mathcal{S}^k$ denotes the closed convex cone of positive semidefinite matrices. We write $A \succ 0$ (resp. $A \succeq 0$) to mean that A is positive definite (resp. positive semidefinite).

2 Triangulation As Polynomial Optimization

With an eye towards the future, let us re-state the triangulation problem (2) as the constrained optimization problem

$$\operatorname*{arg\,min}_{x_1, \ldots, x_n, X} \sum_i \|x_i - \hat{x}_i\|^2, \quad \text{s.t. } x_i = \Pi P_i \tilde{X}, \quad \forall i = 1, \ldots, n. \tag{3}$$

In this formulation, the constraints state that each x_i is the projection of X in image i. Let us now denote the feasible region for this optimization problem by

$$V_P = \left\{ x \in \mathbb{R}^{2n} \mid \exists X \in \mathbb{R}^3 \text{ s.t. } x_i = \Pi P_i \tilde{X},\ \forall i = 1, \ldots, n \right\}. \tag{4}$$

For any $x \in V_P$, we can recover the corresponding X using the SVD based algorithm [4]. Now we can re-state (3) purely in terms of x

$$\arg\min_x \ \|x - \hat{x}\|^2, \quad \text{s.t. } x \in V_P. \tag{5}$$

Let $\overline{V}_P \supseteq V_P$ be the closure of V_P, meaning that $\overline{V}_P$ contains all the limit points of V_P. Then consider the following optimization problem:

$$\arg\min_x \ \|x - \hat{x}\|^2, \quad \text{s.t. } x \in \overline{V}_P. \tag{6}$$

The objective function in these two optimization problems is the squared distance from $\widehat{x}$ to the sets V_P and $\overline{V}_P$ respectively. Since $\overline{V}_P$ is the topological closure of V_P it can be shown that any solution x^* to (6) which is not in V_P is arbitrarily close to a point in V_P, and the optimal objective function values for (5) and (6) are the same. Thus, solving (6) is essentially equivalent to solving (5).

The set V_P is a quasi-projective variety. A *variety* is the zero set of a finite set of polynomials, and a *quasi-projective variety* is the set difference of two varieties. Therefore, $\overline{V}_P$ is also a variety [18]. Heyden & Åström [5] show that

$$\overline{V}_P = \left\{ x \in \mathbb{R}^{2n} \;\middle|\; \begin{array}{l} f_{ij}(x) = 0, \ 1 \leq i < j \leq n \\ t_{i,j,k}(x) = 0, \ 1 \leq i < j < k \leq n \end{array} \right\}. \tag{7}$$

Here $f_{ij}(x) = \tilde{x}_i^\top F_{ij} \tilde{x}_j = 0$ are the bilinear/quadratic epipolar constraints, where $F_{ij} \in \mathbb{R}^{3\times 3}$ is the fundamental matrix for images i and j. The second set of constraints $t_{ijk}(x_i, x_j, x_k) = 0$ are the the trilinear/cubic constraints defined by the trifocal tensor on images i, j and k.

At the risk of a mild abuse of notation, we will also use F_{ij} to denote a $(2n+1) \times (2n+1)$ matrix such that $f_{ij}(x) = \tilde{x}^\top F_{ij} \tilde{x}$. The construction of this matrix involves embedding two copies of the 3×3 fundamental matrix (with suitable reordering of the entries) in an all zero $(2n+1) \times (2n+1)$ matrix.

Now, let W_P be the quadratic variety

$$W_P = \left\{ x \in \mathbb{R}^{2n} \mid \tilde{x}^\top F_{ij} \tilde{x} = 0, \ 1 \leq i < j \leq n \right\}. \tag{8}$$

For $n = 2$, since there are no trilinear constraints $W_P = \overline{V}_P$ but for $n \geq 3$, in general $W_P \supseteq \overline{V}_P$. For $n \geq 4$, Heyden & Åström show that if the camera centers are not co-planar then $W_P = \overline{V}_P$ [5]. Note that for $n = 3$, the camera centers are always co-planar. Therefore, for $n = 2$, and for $n \geq 4$ when the camera centers are non-co-planar, we can just optimize over the quadratic variety W_P:

$$\arg\min_x \ \|x - \hat{x}\|^2 \quad \text{s.t. } x \in W_P. \tag{9}$$

However, we cannot just ignore the co-planar case as a degeneracy since it is a common enough occurrence, e.g., an object rotating on a turntable in front of a fixed camera. If all the camera centers lie on a plane π_P, then solving (9) instead of (6) can result in spurious solutions, i.e. a projection vector x^* for which there is no single point $X^* \in \mathbb{R}^3$ that projects to x_i^* for each image i. This can happen if each x_i^* lies on the image of the plane π_P in image i. It is easy to reject such a spurious x^* by checking if it satisfies the trilinear constraints.

From here on, we will focus our attention on solving (9) and in Section 5 we will show that solving (9) instead of (6) is not a significant source of failures.

Let us now define the polynomial

$$g(x) = \|x - \widehat{x}\|^2 = \tilde{x}^\top \begin{bmatrix} I & -\widehat{x} \\ -\widehat{x}^\top & \|\widehat{x}\|^2 \end{bmatrix} \tilde{x} = \tilde{x}^\top G \tilde{x}, \tag{10}$$

where I is the $2n \times 2n$ identity matrix. Observe that $G \in \mathcal{S}_+^{2n+1}$, and the Hessian of g is $\nabla^2 g = 2I$. Similarly, let $F_{ij} = \begin{bmatrix} H_{ij} & b_{ij} \\ b_{ij}^\top & \beta_{ij} \end{bmatrix}$, where $H_{ij} \in \mathcal{S}^{2n}$, $b_{ij} \in \mathbb{R}^{2n}$, and $\beta_{ij} \in \mathbb{R}$. Then $\nabla^2 f_{ij} = 2H_{ij}$. We re-write (9) as the quadratically constrained quadratic program (QCQP)

$$\arg\min_{x \in \mathbb{R}^{2n}} \tilde{x}^\top G \tilde{x} \quad \text{s.t.} \quad \tilde{x}^\top F_{ij} \tilde{x} = 0,\ 1 \le i < j \le n. \tag{11}$$

3 Semidefinite Relaxation

We re-write (11) as the following rank constrained semidefinite program (SDP).

$$\begin{aligned} \arg\min_{Y}\ & \langle G, Y \rangle \\ \text{s.t.}\ & \langle F_{ij}, Y \rangle = 0,\ 1 \le i < j \le n, \\ & \langle E, Y \rangle = 1, \\ & Y \in \mathcal{S}_+^{2n+1}, \\ & \operatorname{rank}(Y) = 1. \end{aligned} \tag{12}$$

Here $E \in \mathcal{S}^{2n+1}$ is an all zero matrix except for its bottom right entry which equals one. The problems (11) and (12) are equivalent: x is feasible (optimal) for (11) if and only if $Y = \tilde{x}\tilde{x}^\top$ is feasible (optimal) for (12).

Solving rank constrained semidefinite programs is NP-hard [22]. Dropping the rank constraint gives the primal semidefinite program

$$\begin{aligned} \arg\min_{Y}\ & \langle G, Y \rangle \\ \text{s.t.}\ & \langle F_{ij}, Y \rangle = 0,\ 1 \le i < j \le n, \\ & \langle E, Y \rangle = 1, \\ & Y \in \mathcal{S}_+^{2n+1}. \end{aligned} \tag{13}$$

The dual of this primal semidefinite program is

$$\begin{aligned} \arg\max_{\lambda_{ij}, \rho}\ & \rho \\ \text{subject to}\ & G + \sum \lambda_{ij} F_{ij} - \rho E \succeq 0, \\ & \lambda_{ij}, \rho \in \mathbb{R},\ 1 \le i < j \le n. \end{aligned} \tag{14}$$

The primal SDP (13) is also known as the first *moment* relaxation and its dual SDP (14) is known as the first *sum of squares* relaxation. They are instances of a general hierarchy of semidefinite relaxations for polynomial optimization problems [10]. Problem (14) is also the Lagrangian dual of (11).

The remainder of this paper is dedicated to the possibility that solving the triangulation problem is equivalent to solving these semidefinite relaxations. Let us denote by $g^*, g^{\text{quad}}, g^{\text{mom}}, g^{\text{sos}}$, the optimal solutions to the optimization problems (6), (11), (13), (14) respectively. Then the following lemmas hold.

Lemma 1. *For all n, $g^* \geq g^{\text{quad}}$. For $n = 2$, or $n \geq 4$ with non-co-planar cameras, $g^* = g^{\text{quad}}$.*

Proof. The claim follows from the discussion in Section 2 after the definition of W_P. □

Lemma 2. $g^{\text{quad}} \geq g^{\text{mom}}$.

Proof. This is true because (13) is a relaxation of (11). □

Lemma 3. *For all n, $g^{\text{mom}} = g^{\text{sos}}$. Moreover, there exist optimal Y^*, λ^*_{ij} and ρ^* that achieve these values.*

Proof. The inequality $g^{\text{mom}} \geq g^{\text{sos}}$ follows from weak duality. Equality, and the existence of Y^*, λ^*_{ij} and ρ^* which attain the optimal values follow if we can show that the feasible regions of both the primal and dual problems have nonempty interiors [22, Theorem 3.1] (also known as Slater's constraint qualification).

For the primal problem, let $x \in \mathbb{R}^{2n}$ be any feasible point for the triangulation problem (9) (such a feasible point always exists) and let $D = \text{diag}(1, \ldots, 1, 0) \in \mathcal{S}^{2n+1}$. It is easy to show that $Y = \tilde{x}\tilde{x}^\top + D$ is positive definite and primal feasible. For the dual problem, take $\lambda_{ij} = 0$ and $\rho = -1$ and verify $G + E \succ 0$. □

4 The Algorithm and its Analysis

We propose Algorithm 1 as a method for triangulation.

4.1 Correctness

Theorem 1. *Algorithm 1 terminates in time polynomial in n.*

Proof. The proof is based on three facts. One, the primal (13) and dual (14) have descriptions of size polynomial in n, the number of images. Two, SDPs can be solved to arbitrary precision in time which is polynomial in the size of their descriptions [22]. Three, the eigenvalue decomposition of a matrix can be computed in polynomial time. □

Before moving forward, we need the following definition.

Definition 1. *A triangulation problem is* SDP-EXACT *if $g^{\text{quad}} = g^{\text{sos}} = g^{\text{mom}}$, i.e., the relaxations are tight.*

Algorithm 1 Triangulation

```
Require: Image observation vector $\widehat{x} \in \mathbb{R}^{2n}$ and the set of cameras $P = \{P_1, \ldots, P_n\}$.
Solve the primal (13) and dual (14) SDPs to optimal $Y^*, \lambda_{ij}^*$ and $\rho^*$.
$x = Y^*[1 : 2n, 2n+1]$ (i.e., $x$ is the last column of $Y^*$ without its last entry)
Use the SVD algorithm to determine a world point $X \in \mathbb{R}^3$ from $x$.
if $I + \sum \lambda_{ij}^* H_{ij} \succ 0$ then
    if $n = 2$, or $n \geq 4$ and the cameras $P$ are non-co-planar, then
        Return (OPTIMAL, $X$).
    end if
    if $x_i = \Pi P_i \tilde{X}$ $\forall i = 1, \ldots, n$ then
        Return (OPTIMAL, $X$)
    end if
end if
Return (SUBOPTIMAL, $X$).
```

We will first describe the conditions under which a triangulation problem is SDP-EXACT. We will then show that if Algorithm 1 returns OPTIMAL then triangulation is SDP-EXACT, and further, the solution X returned by the algorithm is indeed optimal for triangulation.

Theorem 2. *Let x^* be an optimal solution to the quadratic program* (11)*. The triangulation problem is* SDP-EXACT *if and only if there exist $\lambda_{ij} \in \mathbb{R}$ such that*

$$\text{(i) } \nabla g(x^*) + \sum \lambda_{ij} \nabla f_{ij}(x^*) = 0 \qquad \textit{and} \qquad \text{(ii) } I + \sum \lambda_{ij} H_{ij} \succeq 0. \tag{15}$$

Before proving this theorem, we observe that it is not immediately useful from a computational perspective. Indeed, *a priori* verifying condition (i) requires knowledge of the optimal solution x^*. However, the theorem will help us understand *why* the triangulation problem is so often SDP-EXACT in Section 4.3.

Let $L(x, \lambda_{ij}, \rho) = g(x) + \sum_{ij} \lambda_{ij} f_{ij}(x) - \rho = \tilde{x}^\top (G + \sum \lambda_{ij} F_{ij} - \rho E) \tilde{x}$. Observe that $\nabla_x L(x, \lambda_{ij}, \rho) = \nabla g(x) + \sum \lambda_{ij} \nabla f_{ij}(x)$ and $\nabla_x^2 L(x, \lambda_{ij}, \rho) = 2(I + \sum \lambda_{ij} H_{ij})$. We require that the following two simple lemmas, the proofs of which can be found in the Appendix.

Lemma 4. *If x^* is the optimal solution to* (11) *and λ_{ij} satisfy condition (i), then $L(x, \lambda_{ij}, g(x^*)) = (x - x^*)^\top (I + \sum \lambda_{ij} H_{ij}) (x - x^*)$. Further, if condition (ii) is satisfied as well, then $L(x, \lambda_{ij}, g(x^*)) \geq 0,\ \forall x \in \mathbb{R}^{2n}$.*

Lemma 5. *If $x^\top Ax + 2b^\top x + c \geq 0,\ \forall x$, then $\begin{bmatrix} A & b \\ b^\top & c \end{bmatrix} \succeq 0$.*

Proof (Theorem 2). For the *if* direction, let λ_{ij} satisfy conditions (i) and (ii). Then from Lemma 4 we have $L(x, \lambda_{ij}, g(x^*)) \geq 0\ \forall x \in \mathbb{R}^{2n}$, which combined with Lemma 5 gives $G + \sum \lambda_{ij} F_{ij} - g(x^*)E \succeq 0$. Therefore λ_{ij} and $\rho = g(x^*)$ are dual feasible, which in turn means that $g^{\text{sos}} \geq \rho = g(x^*) = g^{\text{quad}}$. Lemmas 2 and 3 give the reverse inequality, thus $g^{\text{sos}} = g^{\text{mom}} = g^{\text{quad}} = g(x^*)$.

For the *only if* direction, let ρ^* and λ^*_{ij} be the optimal solution to the dual (14). If the problem is SDP-EXACT we have the equality $\rho^* = g(x^*) = g^{\text{quad}}$ and from the dual feasibility of λ^*_{ij} and ρ^* we have that $G + \sum \lambda^*_{ij} F_{ij} - \rho^* E \succeq 0$. Taken together, these two facts imply that $L(x, \lambda^*_{ij}, g(x^*)) \geq 0,\ \forall x \in \mathbb{R}^{2n}$.

Since $L(x^*, \lambda^*_{ij}, g(x^*)) = 0$, $L(x, \lambda^*_{ij}, g(x^*))$ is a non-negative quadratic polynomial that vanishes at x^*. Non-negativity implies condition (ii) (that the Hessian of the polynomial is positive semidefinite) and the fact that zero is the minimum possible value of a non-negative polynomial implies that its gradient vanishes at x^* which is exactly condition (i). □

Condition (ii) of Theorem 2 is automatically satisfied by *any* feasible λ_{ij} for the dual (14). Hence, verifying SDP exactness using the dual optimal λ^*_{ij} reduces to checking condition (i) of Theorem 2. Checking this condition is not computationally practical, since it requires knowledge of the optimum x^*. By slightly tightening condition (ii) we can bypass condition (i).

Theorem 3. *If* $\{Y^*,\ \lambda^*_{ij},\ \rho^*\}$ *are primal-dual optimal and* $I + \sum \lambda^*_{ij} H_{ij} \succ 0$, *then* $\operatorname{rank}(Y^*) = 1$, *and triangulation is* SDP-EXACT.

Proof. Notice that $I + \sum \lambda^*_{ij} H_{ij}$ is the top left $(2n) \times (2n)$ block in the larger $(2n+1) \times (2n+1)$ positive semidefinite matrix $G + \sum \lambda^*_{ij} F_{ij} - \rho^* E$. By hypothesis, $I + \sum \lambda^*_{ij} H_{ij}$ is nonsingular and thus has full rank equal to $2n$, which implies

$$\operatorname{rank}\left(G + \sum \lambda^*_{ij} F_{ij} - \rho^* E\right) \geq 2n. \tag{16}$$

The dual and the primal SDP solutions satisfy *complementary slackness*, which means that $\left\langle G + \sum \lambda^*_{ij} F_{ij} - \rho^* E, Y^* \right\rangle = 0$. In particular it implies that

$$\operatorname{rank}(G + \sum \lambda^*_{ij} F_{ij} - \rho^* E) + \operatorname{rank}(Y^*) \leq 2n + 1, \tag{17}$$

where we use the standard fact that whenever $\langle A, B \rangle = 0$ for $A, B \in \mathcal{S}^N_+$, then $\operatorname{rank}(A) + \operatorname{rank}(B) \leq N$. From (16) and (17) we have $\operatorname{rank}(Y^*) \leq 1$. Since $\langle E, Y \rangle = 1$, we have $Y^* \neq 0$ and hence $\operatorname{rank}(Y^*) = 1$. □

Line 4 of Algorithm 1 uses Theorem 3 to establish that we have solved (11). Lines 5–10 of the algorithm are then devoted to making sure that the solution actually lies in $\overline{V}_P$. Thus, Algorithm 1 is correct.

4.2 Implications

Theorem 4. *If the image observations are noise free, i.e. there exists* $X^* \in \mathbb{R}^3$ *such that* $\widehat{x}_i = \Pi P_i X^*,\ \forall i = 1, \ldots, n$, *then Algorithm 1 returns* OPTIMAL

Proof. Setting $\lambda_{ij} = 0$ satisfies the hypothesis for Theorems 2 and 3. □

Theorem 5. *Two view triangulation is* SDP-EXACT.

Proof. For $n = 2$ the triangulation problem (11) involves minimizing a quadratic objective over a single quadratic equality constraint $f_{12}(x) = 0$. The conditions of Theorem 2 in this case reduce to finding $\lambda \in \mathbb{R}$ satisfying

$$\nabla g(x^*) + \lambda \nabla f_{12}(x^*) = 0 \qquad \text{and} \qquad I + \lambda H_{12} \succeq 0. \tag{18}$$

The existence of such a λ follows directly from [14, Theorem 3.2]. □

Theorems 3 and 5 nearly imply that two-view triangulation can be solved in polynomial time. We say *nearly* because, despite Theorem 5, it is possible that the matrix $I + \lambda^* H_{12}$ is singular for the dual optimal λ^* (see Appendix for an example). This is not a contradiction, since Theorem 3 is only a sufficient condition for optimality. Despite such pathologies, we shall see in Section 5 that in practice Algorithm 1 usually returns OPTIMAL for two-view triangulation.

4.3 Geometry of the Algorithm

Recall that the optimization problem (11) can be interpreted as determining the closest point x^* to $\widehat{x}$ in the variety W_P. This viewpoint gives geometric intuition for why Algorithm 1 can be expected to perform well in practice.

Lemma 6. *Given $\widehat{x}$, and assuming appropriate regularity conditions at the optimal solution x^* to* (11)*, there exist $\lambda_{ij} \in \mathbb{R}$ such that*

$$\widehat{x} = x^* + \sum \frac{\lambda_{ij}}{2} \nabla f_{ij}(x^*) \tag{19}$$

Proof. It follows from Lagrange multiplier theory [15] that there exist Lagrange multipliers $\lambda_{ij} \in \mathbb{R}$ such that $\nabla g(x^*) + \sum \lambda_{ij} \nabla f_{ij}(x^*) = 0$. Observe that for (11) $\nabla g(x^*) = 2(x^* - \widehat{x})$, which finishes the proof. □

If $\|x^* - \widehat{x}\|$ is small, i.e. $\widehat{x}$ is close to the variety W_P, then there must exist some λ_{ij} satisfying (19) such that $\|\lambda\|$ is small and hence $I + \sum \lambda_{ij} H_{ij}$ is a small perturbation of the positive definite identity matrix I. Since I lies in the interior of the positive semidefinite cone, these small perturbations also lie in the interior, that is $I + \sum \lambda_{ij} H_{ij} \succ 0$.

This, coupled with the fact that λ_{ij} are Lagrange multipliers at x^*, yields the sufficient conditions in Theorem 2 for a triangulation problem to be SDP-EXACT. Thus, if the amount of noise in the observations is small, Algorithm 1 can be expected to recover the optimal solution to the triangulation problem. Since H_{ij} depends only on the cameras P_i and not on $\widehat{x}$, the amount of noise which Algorithm 1 will tolerate depends only on P_i. We summarize this formally in the following theorem.

Theorem 6. *Let $N(x^*) = \{x^* + \sum \frac{\lambda_{ij}}{2} \nabla f_{ij}(x^*) \mid I + \sum \lambda_{ij} H_{ij} \succ 0\}$. For any $\widehat{x} \in \mathbb{R}^{2n}$, if Algorithm 1 returns* OPTIMAL *and x^* is the optimal image projection vector then $\widehat{x} \in N(x^*)$. Conversely, if $\widehat{x} \in N(x^*)$ and x^* is the closest point in W_P to $\widehat{x}$, then Algorithm 1 will return* OPTIMAL*.*

Proof. The proof is a straightforward application of Lemma 6 and Theorem 3. □

5 Experiments

Algorithm 1 was implemented using YALMIP [12], SeDuMi [21] and MATLAB. These tools allow for easy implementation and the timings below are for completeness and should not be used to judge the runtime performance of the algorithm.

Fundamental matrices and epipolar constraints are specified only up to scale and badly scaled fundamental matrices lead to poorly conditioned SDPs. This was easily fixed by dividing each F_{ij} by its largest singular value.

Algorithm 1 verifies optimality by checking the positive definiteness of $I + \sum \lambda^*_{ij} H_{ij} \succ 0$, which requires that its smallest eigenvalue be greater than some $\delta > 0$. In all our experiments we set $\delta = 0.05$.

Either Algorithm 1 returns OPTIMAL and the solution is guaranteed to be optimal, or it returns SUBOPTIMAL in which case we cannot say anything about the quality of the solution, even though it could still be optimal or at least serve as a good guess for non-linear iteration. Thus, we run Algorithm 1 on a number of synthetic and real-world datasets and report the fraction of cases in which the algorithm certifies optimality.

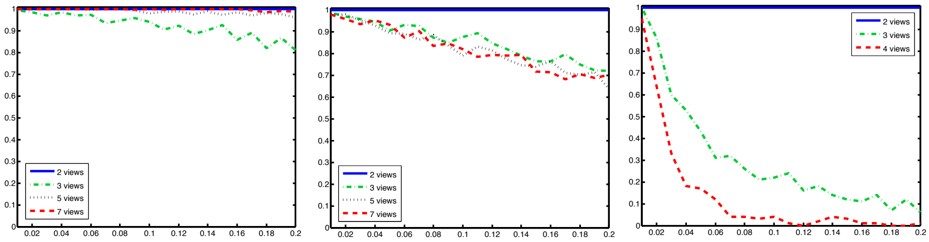
5.1 Synthetic Data

To test the performance of the algorithm as a function of camera configuration and image noise we first tested Algorithm 1 on three synthetic data sets. Following [16], we created instances of the triangulation problem by randomly generating points inside the unit cube in $\mathbb{R}^3$.

For the first experiment, a varying number of cameras (2, 3, 5 and 7) were placed uniformly at random on a sphere of radius 2. In the second experiment, the same number of cameras were uniformly distributed on a circle of radius 2 in the xy-plane. In the third experiment they were restricted to the x-axis and were placed at a distance of 3, 5, 7, and 9 units. These setups result in image measurements with an approximate square length of 2 units. Gaussian noise of varying standard deviation was added to the image measurements. The maximum standard deviation was 0.2, which corresponds to about 10% of the image size. For each noise level we ran the experiment 375 times. Figures 5.1(a), (b) and (c) show the fraction of test cases for which Algorithm 1 returned OPTIMAL as a function of the the standard deviation of perturbation noise.

We first note that for $n = 2$ cameras, in all three cases, independent of camera geometry and noise, we are able to solve the triangulation problem optimally 100% of the time. This experimentally validates Theorem 5 and provides strong evidence that for $n = 2$ there is no gap between Theorems 2 and 3 in practice.

Observe that for cameras on a sphere (Figure 5.1(a)), the algorithm performs very well, and while the performance drops as the noise is increased, the drop is not significant for practical noise levels. Another interesting feature of this graph is that finding and certifying optimality becomes easier with increasing number of cameras. The reason for this increase in robustness is not clear and we plan to further pursue this intriguing observation in future research.



(a) Cameras on a sphere. (b) Cameras on a circle. (c) Cameras on a line.

Fig. 1. Fraction of synthetic experiments in which Algorithm 1 returned OPTIMAL versus the standard deviation of the noise level added to the images. (a) Cameras placed randomly on the sphere of radius 2. (b) Three cameras all placed on the xy-plane. (c) Three cameras placed along the x-axis.

Cameras on a circle (Figure 5.1(b)) is one of the degenerate cases of our algorithm where $\overline{V}_P \neq W_P$. We expect a higher rate of failure here and indeed this is the case. It is however worth noting that the algorithm does not immediately breakdown and shows a steady decline in performance as a function of noise like the previous experiment. Unlike cameras on a sphere, increasing the number of cameras does not increase the performance of the algorithm, which points to the non-trivial gap between $\overline{V}_P$ and W_P for co-planar cameras.

Finally let us consider the hard problem of triangulating a point when the camera is moving on a line. This case is hard geometrically and algorithmically as the cameras are co-planar, and this difficulty is reflected in the performance of the algorithm. In contrast to the previous experiments, we observe rapid degradation with increasing noise.

5.2 Real Data

We tested Algorithm 1 on four real-world data sets: the Model House, Corridor, and Dinosaur from Oxford University and the Notre Dame data set [19]. Our results are summarized in Table 1.

Table 1. Performance of Algorithm 1 on real data. The column OPTIMAL reports the fraction of triangulation problems which were certified optimal.

Data set	# images	# points	OPTIMAL	Time (sec)
Model House	10	672	1.000	143
Corridor	11	737	0.999	193
Dinosaur	36	4983	1.000	960
Notre Dame	48	16,288	0.984	7200

Model house is a simple data set where the camera moves laterally in front of the model house. Global optimality was achieved in all cases.

Corridor is a geometrically hard sequence where most of the camera motion is forward, straight down a corridor. This is similar to the synthetic experiment where the cameras were all on the x-axis. The algorithm returned OPTIMAL in all but one case.

Dinosaur consists of images of a (fake) dinosaur rotating on a turntable in front of a fixed camera. Even though the camera configuration is hard for our algorithm (cameras are co-planar), global optimality was achieved in all cases.

The three Oxford datasets are custom captures from the same camera. Notre Dame consists of images of the Notre Dame cathedral downloaded from Flickr. The data set comes with estimates of the radial distortion for each camera, which we ignore as we are only considering projective cameras in this paper. Algorithm 1 returned OPTIMAL in 98.36% of cases.

It is worth noting here that the synthetic datasets were designed to test the limits of the algorithm as a function of noise. A standard deviation equal to 0.2 translates to image noise of approximately 10% of the image size. In practice such high levels of noise rarely occur, and when they do, they typically correspond to outliers. The results in Table 1 indicate that the algorithm has excellent performance in practice.

6 Discussion

We have presented a semidefinite programming algorithm for triangulation. In practice it usually returns a global optimum and a certificate of optimality in polynomial time. Of course there are downsides which must be taken into consideration. Solving SDPs is not nearly as fast as gradient descent methods. Moreover, what happens in the rare cases that global optimality is not certifiable? Regardless, the lure of a polynomial time algorithm and a better understanding of the geometry of the triangulation problem is far too great to allow these hiccups to close the door on SDPs.

Hartley and Sturm [3] show that two-view triangulation can be solved by finding the roots of a degree 6 univariate polynomial. For $n = 3$, [20] gives a Gröbner basis based algorithm for triangulation by finding all solutions to a polynomial system with 47 roots. These methods do not extend in a computationally useful manner to $n \geq 4$. Our algorithm solves the triangulation problem for all values of $n \geq 2$ under the conditions of Theorem 3 when the camera centers are not co-planar.

Similar to our setup, Kanatani et al. also frame triangulation as finding the closest point on a variety from noisy observations [9]. Unlike our work, they use both epipolar and trifocal constraints, which makes the constraint set much more complicated. Further, they do not prove finite convergence or the optimality of their procedure. Our work answers their question of analyzing the shape of the variety to obtain a *noise threshold* for optimality guarantees.

Semidefinite relaxations for optimization problems in multi-view geometry were first studied by Kahl & Henrion in [8]. They formulate triangulation as an optimization problem over quartic inequalities and study its moment relaxations. Kahl & Henrion observe that good solutions can be obtained from the first few relaxations. In our method the very first relaxation already yields high quality solutions. Further, the quadratic equality based formulation has nice theoretical properties that explain the empirical performance of our algorithm.

Hartley & Seo proposed a method for verifying global optimality of a given solution to the triangulation problem [2]. The relation between their test and the definiteness test of Theorem 3 is a fascinating direction for future research.

The study of QCQPs has a long history. There is a wide body of work devoted to verifying optimality of a given solution for various classes of QCQPs and using semidefinite programming to solve them. Thus it is natural that statements similar to Theorems 2 and 3 for various subclasses of QCQPs have appeared several times in the past e.g. [6, 11, 17].

The astute reader will notice that even though Theorems 2 and 3 are presented for triangulation, the proofs apply to any QCQP with equality constraints, and can be interpreted in terms of the Karush-Kuhn-Tucker (KKT) conditions for (11). More recently, and independently from us, Zheng et al. have proved versions of Theorems 2 and 3 for the more general case of inequality-constrained QCQPs [23]. Our formulation of triangulation as minimizing distance to a variety allows for a geometric interpretation of these optimality conditions, which in turn explains the performance of our algorithm.

Finally, we use the general purpose SDP solver `SeDuMi` for our experiments. Much improvement can be expected from a custom SDP solver which makes use of the explicit symmetry and sparsity of the triangulation problem (e.g. for all n the linear matrix constraints in the primal semidefinite program (13) are sparse and have rank at most five).

Acknowledgements

It is a pleasure to acknowledge our conversations with Pablo Parrilo, Jim Burke, Greg Blekherman and Fredrik Kahl. In particular Fredrik Kahl introduced us to the work of Heyden & Åström [5]. Chris Aholt and Rekha Thomas were supported by the NSF grant DMS-1115293.

References

1. R. W. Freund and F. Jarre. Solving the sum-of-ratios problem by an interior-point method. *J. Glob. Opt.*, 19(1):83–102, 2001.
2. R. Hartley and Y. Seo. Verifying global minima for l_2 minimization problems. In *CVPR*, 2008.
3. R. Hartley and P. Sturm. Triangulation. *CVIU*, 68(2):146–157, 1997.
4. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, second edition, 2003.

5. A. Heyden and K. Åström. Algebraic properties of multilinear constraints. *Math. Methods Appl. Sci.*, 20(13):1135–1162, 1997.
6. V. Jeyakumar, A. Rubinov, and Z. Wu. Non-convex quadratic minimization problems with quadratic constraints: global optimality conditions. *Mathematical Programming*, 110(3):521–541, 2007.
7. F. Kahl, S. Agarwal, M. K. Chandraker, D. J. Kriegman, and S. Belongie. Practical global optimization for multiview geometry. *IJCV*, 79(3):271–284, 2008.
8. F. Kahl and D. Henrion. Globally optimal estimates for geometric reconstruction problems. *IJCV*, 74(1):3–15, 2007.
9. K. Kanatani, H. Niitsuma, and Y. Sugaya. Optimization without search: Constraint satisfaction by orthogonal projection with applications to multiview triangulation. *Memoirs of the Faculty of Engineering*, 44:32–41, 2010.
10. M. Laurent. Sums of squares, moment matrices and optimization over polynomials. volume 149 of *IMA Vol. Math. Appl.*, pages 157–270. Springer, 2009.
11. G. Li. Global quadratic minimization over bivalent constraints: Necessary and sufficient global optimality condition. *Journal of Optimization Theory and Applications*, pages 1–17, 2012.
12. J. Lofberg. YALMIP: A toolbox for modeling and optimization in matlab. In *Int. Symp. on Computer Aided Control Systems Design*, pages 284–289, 2004.
13. F. Lu and R. I. Hartley. A fast optimal algorithm for l_2 triangulation. In *ACCV*, volume 2, pages 279–288, 2007.
14. J. Moré. Generalizations of the trust region problem. *Optimization methods and Software*, 2(3-4):189–209, 1993.
15. J. Nocedal and S. Wright. *Numerical Optimization.* Springer Verlag, 1999.
16. C. Olsson, F. Kahl, and R. Hartley. Projective least-squares: Global solutions with local optimization. In *CVPR*, pages 1216–1223, 2009.
17. M. Pinar. Sufficient global optimality conditions for bivalent quadratic optimization. *Journal of optimization theory and applications*, 122(2):433–440, 2004.
18. I. Shafarevich. *Basic Algebraic Geometry I: Varieties in Projective Space.* Springer-Verlag, 1998.
19. N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3d. In *SIGGRAPH*, pages 835–846, 2006.
20. H. Stewenius, F. Schaffalitzky, and D. Nister. How hard is 3-view triangulation really? In *ICCV*, pages 686–693, 2005.
21. J. Sturm. Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones. *Opt. Meth. and Soft.*, 11-12:625–653, 1999.
22. L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.
23. X. Zheng, X. Sun, D. Li, and Y. Xu. On zero duality gap in nonconvex quadratic programming problems. *Journal of Global Optimization*, 52(2):229–242, 2012.

Appendix

Failure of Theorem 3 for two-view triangulation

Consider the cameras

$$A_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & a \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & b \end{bmatrix}, \tag{20}$$

where $0 < a < b$. This models the situation in which two cameras are placed on the x-axis at $(a, 0, 0)$ and $(b, 0, 0)$, respectively, both facing the origin (i.e. viewing down the x-axis).

We can show that, given $\varepsilon > 0$ and $\widehat{x} = (\widehat{x}_1, \widehat{x}_2) = \left(\left[0; \varepsilon\right], \left[\varepsilon; 0\right]\right)$, the minimum value of the QCQP (11) is ε^2 and all points of the form

$$x_1^* = \left[\sqrt{\mu(\varepsilon - \mu)};\, \mu\right] \tag{21}$$

$$x_2^* = \left[\varepsilon - \mu;\, \sqrt{\mu(\varepsilon - \mu)}\right] \tag{22}$$

where $0 \leq \mu \leq \varepsilon$ are optimal.

The non-uniqueness of the minimizer implies that the verification matrix $I + \lambda H_{12}$ cannot be positive definite. So in this case, the hypothesis of Theorem 3 is not satisfied, although Theorem 5 still holds.

The proof of these facts is a straight-forward application of the KKT conditions. The minimum distance ε^2 can be attained by setting one of the image points equal to the epipole, but not both. However, the minimum distance is also attained at points which correspond to neither image being the epipole.

Proof of Lemma 4

Recall conditions (i) $\nabla g(x^*) + \sum \lambda_{ij} \nabla f_{ij}(x^*) = 0$, and (ii) $I + \sum \lambda_{ij} H_{ij} \succeq 0$. Let $H = \sum \lambda_{ij} H_{ij}$, $b = \sum \lambda_{ij} b_{ij}$, and $\beta = \sum \lambda_{ij} \beta_{ij}$. Then

$$L(x, \lambda_{ij}, g(x^*)) = g(x) + \sum \lambda_{ij} f_{ij}(x) - g(x^*) \tag{23}$$

$$= x^\top (I + H)\, x + 2\,(b - \widehat{x})^\top x + \widehat{x}^\top \widehat{x} + \beta - g(x^*). \tag{24}$$

We wish to show that (24) is equal to $(x - x^*)^\top (I + H)\,(x - x^*)$. The fact that $L(x, \lambda_{ij}, g(x^*)) \geq 0$ is then immediate because $I + H \succeq 0$. Looking at the quadratic, linear, and constant terms separately, it suffices to prove

$$\text{(a) } (I + H)x^* = \widehat{x} - b \qquad \text{and} \qquad \text{(b) } {x^*}^\top (I + H)\, x^* = \widehat{x}^\top \widehat{x} + \beta - g(x^*). \tag{25}$$

Indeed, (a) is just a restatement of condition (i), since $\nabla g(x) = 2(x - \widehat{x})$ and $\nabla f_{ij}(x) = 2(H_{ij} x + b_{ij})$. To prove (b), first recall that since x^* is feasible for problem (11), $f_{ij}(x^*) = 0$ for all i, j. In particular, $\beta = -2b^\top x^* - {x^*}^\top H x^*$. Using this and part (a), one can verify (b) through straightforward manipulations. □

Proof of Lemma 5

Suppose the matrix $M = \left[A, b\,; b^\top, c\right]$ is not positive semidefinite, i.e. there exists y such that $y^\top M y < 0$. Write $y = [y'; \gamma]$ for some $\gamma \in \mathbb{R}$. If $\gamma = 0$, then $0 > y^\top M y = {y'}^\top A y'$. But then we arrive at a contradiction by considering $x = \mu y'$ for a scalar $\mu \in \mathbb{R}$, since for μ large enough, $x^\top A x + 2b^\top x + c < 0$.

Now if $\gamma \neq 0$, setting $x = y'/\gamma$ gives

$$x^\top A x + 2b^\top x + c = \frac{1}{\gamma^2}\left(y^\top M y\right) < 0, \tag{26}$$

which again contradicts the hypothesis. □