An Efficient Solution to Non-Minimal Case Essential Matrix Estimation

Ji Zhao

Abstract—Finding relative pose between two calibrated views is a fundamental task in computer vision. Given the minimal number 5 of required point correspondences, the classical five-point method can be used to calculate the essential matrix. For the non-minimal cases when N (N > 5) correct point correspondences are given, which is called N-point problem, methods are relatively less mature. In this paper, we solve the N-point problem by minimizing the algebraic error and formulate it as a quadratically constrained quadratic program (QCQP). The formulation is based on a simpler parameterization of the feasible region – the normalized essential matrix manifold – than previous approaches. Then a globally optimal solution to this problem is obtained by semidefinite relaxation. This allows us to obtain certifiably global solutions to an important non-convex problem in polynomial time. We provide the condition to recover the optimal essential matrix from the relaxed problems. The theoretical guarantees of the semidefinite relaxation are investigated, including the tightness and local stability. Experiments demonstrate that our approach always finds and certifies (a-posteriori) the global optimum of the cost function, and it is dozens of times faster than state-of-the-art globally optimal solutions.

Index Terms—Relative pose estimation, ego-motion estimation, essential matrix manifold, epipolar geometry, quadratically constrained quadratic program, semidefinite programming, convex optimization

1 INTRODUCTION

Finding relative pose between two calibrated views using 2D-2D correspondences is a cornerstone in geometric vision [1], and it makes a basic building block in many structure-from-motion (SfM), visual odometry, and simultaneous localization and mapping (SLAM) systems. The essential matrix for calibrated cameras, which encodes the relative pose in two projective views, is the most popular representation of relative pose [2].

Due to scale ambiguity of translation, the relative pose has only 5 degrees-of-freedom (DoFs), including 3 for rotation and 2 for translation. Except for degenerate configurations, 5 point correspondences are hence sufficient to determine the relative pose. Given 5 point correspondences, the five-point method using essential matrix representation [1] can be used to calculate the essential matrix efficiently. The above solution is the so-called minimal solution to the problem. The minimal problem is usually integrated into a hypothesis-and-test framework, such as RANSAC [3], to find the solution corresponds to the largest consensus set. Hence it gains robustness against outlier correspondences. Once the maximal consensus set is found in such a framework, it is required that all inliers are used to re-estimate a model to reduce the influence of noise. In relative pose estimation context, this is known as non-minimal relative pose estimation. We will call it N-point problem in this paper. As it has been investigated in [4], [5], [6], non-minimal relative pose estimation will lead to more accurate result than minimal case solutions.

Though it seems to be a rather fundamental problem, it is surprising that N-point problem receives very little attention. The well-known direct linear transformation (DLT) technique [2] can be used to estimate the essential matrix using 8 or more point correspondences. However, DLT ignores the inherent nonlinear constraints of the essential matrix. To remedy this problem, an essential matrix is recovered after obtaining an approximated essential matrix from the DLT solution. Recently an eigenvalue-based formulation and its variant were proposed to solve the N-point problem [5], [6]. However, all above mentioned methods fail to guarantee the global optimality and efficiency simultaneously.

The progress of N-point problem is far behind the absolute pose estimation from 2D-3D point correspondences (perspective-n-point (PnP)) [7] and 3D-3D point correspondences [8], [9]. Take PnP as an example, the EPnP [7] is rather efficient and has linear complexity with the number of observations. It has been successfully used in RANSAC framework given an arbitrary number of inlier 2D-3D correspondences. Arguably, relative pose estimation is more difficult than absolute pose estimation. It is desirable to find a practical solution to N-point problem, whose efficiency and global optimality are both satisfactory. This is the motivation of this paper.

Among the existing parameterizations for the relative pose, the essential matrix is the most popular one. Despite its great success, the essential matrix parameterization suffers from the pure rotation and solution multiplicity. To overcome these problems, some methods that use rotation matrix parameterization have been proposed [5], [6]. More recent advances have shown that the rotation matrix can still be accurately recovered from the essential matrix for pure rotations situations [10]. Besides, some criteria have been proposed to detect pure rotation, and the unique correct solution can be selected without triangulation. Based on this progress, this paper also utilizes essential matrix parameterization due to its simplicity.

The paper is organized as follows. Section 2 describes the background materials and related works. In Section 3

• J. Zhao is with TuSimple, Beijing 100020, China. E-mail: zhaoji84@gmail.com
we introduce a novel parameterization of essential matrix manifold and provide formulations of minimizing the algebraic error. Based on this, Section 4 derives a convex optimization approach by semidefinite relaxation (SDR). Section 5 provides the proof for tightness and local stability of SDR. Section 6 outlines the performance of our method in comparison to other approaches, followed by a concluding discussion in Section 7.

2 Related Works

Finding optimal essential matrix by $L_\infty$ norm objective and branch-and-bound method was proposed in [11]. An algebraic error for essential matrix estimation was investigated in [4]. Its global optimality was obtained by a square matrix representation of homogeneous forms and relaxation. An eigenvalue-based formulation was proposed to estimate rotation matrix [5], in which the problem is optimized by local gradient descent or branch-and-bound. It was later improved by a certifiably globally optimal solution by relaxation and semidefinite programming (SDP) [6].

Another related area is non-minimal fundamental matrix estimation for uncalibrated cameras. Fundamental matrix has fewer constraints than the essential matrix. A matrix is a fundamental matrix if and only if it has two non-zero singular values. An essential matrix has the additional property that the two nonzero singular values are equal [2]. Algebraic error has been widely used in fundamental matrix estimation [12], [13], [14]. In [13], a method for minimizing the algebraic error is proposed which ensures the rank constraint. However, it does not guarantee to avoid local minima in the general case. In [15], [16], the constraint for a fundamental matrix is imposed by setting its determinant to 0, leading to a cubic polynomial constraint. In [14], the fundamental matrix estimation problem is reduced to one or several constrained polynomial optimization problems by imposing the constraint that the null space of the solution must contain a non-zero vector.

For both the essential matrix and fundamental matrix, pose estimation from $N$-point using algebraic error can be formulated as polynomial optimization problems [17]. Polynomial optimization problems can always be rewritten as quadratically constrained quadratic programs (QCQPs), which has a broad of methods to solve. In multiple view geometry, the semidefinite relaxation (SDR) for polynomial optimization problems was first studied by Kahl & Henrion in [18]. Recently, a number of certifiably correct solutions to estimation problems in computer vision and robotics have been developed. For example, SDR and Lagrangian duality of QCQPs have been used in point cloud registration [19], triangulation [20], pose graph optimization [21], and special Euclidean group synchronization [22]. These methods either rely on SDR or the dual form of the primal problem. However, SDR does not guarantee a priori that this approach generates an optimal solution. The global optimality for general QCQPs is still an open problem. When the QCQP satisfies certain conditions and data noise lies within a critical threshold, a recent study proves that the solution to the SDP optimization algorithm is guaranteed to be globally optimal [23], [24].

Essential matrix or fundamental matrix estimation by algebraic error has been extensively exploited in previous literature [4], [5], [6], [12], [13], [14]. Perhaps the most related paper to our paper is the work of [6], which converts the $N$-point problem of an eigenvalue-based formulation to a QCQP by SDR. However, its efficiency is not satisfactory and the tightness of SDR has not been proved. This present paper minimizes an equivalent algebraic error and reformulates it as a much simpler QCQP. The proofs of tightness and local stability are provided which are based on recent works of Cifuentes et al. [23], [24]. The contribution of this paper is three-fold.

- **Formulation.** A novel parameterization is proposed to characterize the normalized essential matrix manifold, which has fewer constraints than others. An $N$-point formulation is proposed based on this parameterization.

- **Efficiency.** Based on the proposed parameterization, we obtained a formulation for $N$-point problem which has much fewer variables and constraints. Its solver is over 80 times faster than state-of-the-art method. The optimality conditions to recover the solution for the original QCQP is provided.

- **Theoretical aspects.** We provide the theoretical guarantees of the SDR relaxation, including SDR tightness and local stability with observation noise. To the best of my knowledge, this is the first time to achieve such kinds of results for $N$-point problem.

3 Formulations of Minimizing the Algebraic Error

In this work, we assume to be in the central and calibrated camera. Denote $(\mathbf{p}_i, \mathbf{p}_i')$ is the $i$-th point correspondence of the same 3D world point from two distinct viewpoints. They are represented as homogeneous coordinates in normalized image plane $\mathbb{F}_n$. Each point in the normalized image plane can be translated into a unique unit bearing vector originating from the camera center. Let $(\mathbf{f}_i, \mathbf{f}_i')$ denotes a correspondence of bearing vectors pointing at the same 3D world point from two distinct viewpoints, where $\mathbf{f}_i$ represents the observation from the first viewpoint, and $\mathbf{f}_i'$ the one from the second. The bearing vectors are determined by $\mathbf{f}_i = \frac{\mathbf{p}_i}{\left\| \mathbf{p}_i \right\|}$ and $\mathbf{f}_i' = \frac{\mathbf{p}_i'}{\left\| \mathbf{p}_i' \right\|}$.

The relative pose is given by the translation $t$ – expressed in the first frame and denoting the position of the second frame w.r.t. the first one – and the rotation $R$ – transforming vectors from the second into the first frame. The normalized direction $t = [t_1, t_2, t_3]^\top$ will be identified with points in the 2-sphere $S^2$, i.e.,

$$S^2 \triangleq \{ t \in \mathbb{R}^3 | t^\top t = 1 \}.$$  

The 3D rotation will be featured as $3 \times 3$ orthogonal matrix with positive determinant belonging to the Special Orthogonal group $SO(3)$, i.e.,

$$SO(3) \triangleq \{ R \in \mathbb{R}^{3 \times 3} | R^\top R = I, \det(R) = 1 \},$$  

where $I$ is an identity matrix.

1. Bold capital letters denote matrices (e.g., $E$ and $R$); bold lower-case letters denote column vectors (e.g., $e, t$); non-bold lower-case letters represent scalars (e.g., $\lambda$). Vectors whose entries are separated by semicolon/comma stands for column/row vectors.
3.1 Essential Matrix Manifold

The essential matrix \( \mathbf{E} \) is defined as \[^2\]
\[
\mathbf{E} = [\mathbf{t}] \times \mathbf{R},
\]
(3)
where subscript \( \times \) constructs the corresponding skew-symmetric matrix for a 3-dimensional vector, i.e.,
\[
[t]_\times = \begin{bmatrix}
t_1 \\
t_2 \\
t_3
\end{bmatrix} = \begin{bmatrix}
0 & -t_3 & t_2 \\
t_3 & 0 & -t_1 \\
-t_2 & t_1 & 0
\end{bmatrix}.
\]
(4)

Denote the essential matrix \( \mathbf{E} \) as
\[
\mathbf{E} = \begin{bmatrix}
e_1 \\
e_2 \\
e_3
\end{bmatrix} = \begin{bmatrix}
e_{11} & e_{12} & e_{13} \\
e_{21} & e_{22} & e_{23} \\
e_{31} & e_{32} & e_{33}
\end{bmatrix},
\]
(5)
where \( e_i \) is the \( i \)-th row of \( \mathbf{E} \). Denote its corresponding vector as
\[
e \triangleq \text{vec}(\mathbf{E}) = [e_{11}, e_{21}, e_{31}, e_{12}, e_{22}, e_{32}, e_{13}, e_{23}, e_{33}]^T,
\]
(6)
where \( \text{vec}(\cdot) \) means stacking all the entries of a matrix by column-first order.

In this paper, an essential matrix set is defined as
\[
\mathcal{M}_E \triangleq \{ \mathbf{E} \mid \mathbf{E} = [\mathbf{t}] \times \mathbf{R}, \exists \mathbf{R} \in \text{SO}(3), \mathbf{t} \in S^2 \}.
\]
(7)
This essential matrix set is called normalized essential matrix manifold \[^{25},^{26}\]. Theorem \[^1\] provides an equivalent condition to define \( \mathcal{M}_E \), which will greatly simplify the optimization in our method.

Theorem 1. A real \( 3 \times 3 \) matrix, \( \mathbf{E} \), is an element in \( \mathcal{M}_E \) if and only if there exists a vector \( \mathbf{t} \in \mathbb{R}^3 \) satisfying the following two conditions:

(i) \( \mathbf{EE}^T = [\mathbf{t}]_\times [\mathbf{t}]_\times^T \) and (ii) \( \mathbf{t}^T \mathbf{t} = 1 \).

Proof. For \( \mathbf{t} \) direction, first it can be verified that
\[
\det ([\mathbf{t}]_\times [\mathbf{t}]_\times - \sigma \mathbf{I}) = -\sigma (\sigma - (t_1^2 + t_2^2 + t_3^2))^2 = -\sigma (\sigma - 1)^2.
\]
Combining this result with condition (i), we can see \( \mathbf{EE}^T \) has an eigenvalue 1 with multiplicity 2 and an eigenvalue 0. According to the definition of singular value, the nonzero singular values of \( \mathbf{E} \) are the square roots of the nonzero eigenvalues of \( \mathbf{EE}^T \). Thus the two nonzero singular values of \( \mathbf{E} \) are equal to 1. According to Theorem 1 in \[^{27}\], \( \mathbf{E} \) is an essential matrix. By combining condition (ii), \( \mathbf{E} \) is an element in \( \mathcal{M}_E \).

For \( \mathbf{t} \) direction, \( \mathbf{E} \) is supposed to be an essential matrix \( \mathcal{M}_E \). According to the definition of \( \mathcal{M}_E \), there exists a vector \( \mathbf{t} \) satisfying condition (ii). Besides, there exists a rotation matrix \( \mathbf{R} \) such that \( \mathbf{E} = [\mathbf{t}]_\times \mathbf{R} \). It can be verified that \( \mathbf{EE}^T = ([\mathbf{t}]_\times \mathbf{R}) ([\mathbf{t}]_\times \mathbf{R})^T = [\mathbf{t}]_\times [\mathbf{t}]_\times^T \), thus condition (i) is also satisfied.

It should be mentioned that a proposition for general essential matrix, which is similar to the \( \text{only if} \) direction in Theorem \[^1\] was presented as Proposition 2 in \[^{27}\] and also Lemma 7.2 in \[^{28}\].

3.2 Minimizing the Algebraic Error

For noise-free cases, the epipolar constraint implies that \[^2\]
\[
f_i^T \mathbf{E} f_i' = 0.
\]
(9)
Due to the existence of noise, this equality will not strictly hold. Instead we pursue the optimal pose by minimizing an algebraic error
\[
\min_{\mathbf{E} \in \mathcal{M}_E} \sum_{i=1}^{N} (f_i^T \mathbf{E} f_i')^2.
\]
(10)
In this problem, the objective is called algebraic error and has been widely used in previous literature \[^4\], \[^{12}\], \[^{13}\], \[^{14}\].

The objective in problem (10) can be reformulated as a quadratic form
\[
\sum_{i=1}^{N} (f_i^T \mathbf{E} f_i')^2 = \mathbf{e}^T \mathbf{C} \mathbf{e},
\]
(11)
where
\[
\mathbf{C}_{9 \times 9} = \sum_{i=1}^{N} (f_i' \otimes f_i) (f_i' \otimes f_i)^T,
\]
(12)
and \( \otimes \) means Kronecker product. Note that \( \mathbf{C} \) is a Gram matrix, so it is positive definite and symmetric.

3.3 QCQP Formulations

By explicitly writing the constraints for the essential matrix manifold \( \mathcal{M}_E \), the problem (10) of minimizing algebraic error can be reformulated as
\[
\min_{\mathbf{E}, \mathbf{R}, \mathbf{t}} \mathbf{e}^T \mathbf{C} \mathbf{e}
\]
\[
s.t. \quad \mathbf{E} = [\mathbf{t}]_\times \mathbf{R}, \ \mathbf{R} \in \text{SO}(3), \ \mathbf{t} \in S^2,
\]
(13)
This problem is a QCQP: The objective is sum-of-squares (SoS), which is positive semidefinite quadratic polynomials; the constraint on the translation vector, \( \mathbf{t}^T \mathbf{t} = 1 \), is also quadratic; a rotation matrix \( \mathbf{R} \) can be fully defined by 20 quadratic constraints \[^{29}\] and, lastly, the relationship between \( \mathbf{E}, \mathbf{R} \) and \( \mathbf{t}, \mathbf{E} = [\mathbf{t}]_\times \mathbf{R} \), is also quadratic. This formulation has 21 unknowns and 30 constraints.

According to Theorem \[^1\], an equivalent QCQP form of minimizing the algebraic error is
\[
\min_{\mathbf{E}, \mathbf{t}} \mathbf{e}^T \mathbf{C} \mathbf{e}
\]
\[
s.t. \quad \mathbf{EE}^T = [\mathbf{t}]_\times [\mathbf{t}]_\times^T, \quad \mathbf{t}^T \mathbf{t} = 1
\]
(14)
There are 12 variables and 7 constraints in this problem. The constraints can be written explicitly as below
\[
\begin{align*}
\begin{bmatrix}
h_1 \\
h_2 \\
h_3 \\
h_4 \\
h_5 \\
h_6 \\
h_7 
\end{bmatrix} &= \begin{bmatrix}
e_1^T e_1 - (t_2^2 + t_3^2) = 0 \\
e_2^T e_2 - (t_1^2 + t_3^2) = 0 \\
e_3^T e_3 - (t_1^2 + t_2^2) = 0 \\
e_1^T e_2 + t_1 t_2 = 0 \\
e_1^T e_3 + t_1 t_3 = 0 \\
e_2^T e_3 + t_2 t_3 = 0 \\
(t_1 t_2 t_3 - 1) = 0
\end{bmatrix}.
\end{align*}
\]
Problems (13) and (14) are equivalent because the objectives are the same and their feasible regions are equivalent. Both of them are nonconvex. In Appendix A, it provides another equivalent optimization problem (35). In the following text, we will only consider problem (14) due to its simplicity. It is homogeneous and does not need any homogenization, which is also simpler than the alternative formulations.

Remark: Both the problem formulations (13) and (14) are equivalent to an eigenvalue-based formulation (9). In the supplementary material of [6], it provides a proof that the objective in eigenvalue-based formulation is equivalent to the algebraic error. Since all formulations essentially utilize the normalized essential matrix manifold as feasible regions and have the equivalent objectives, these formulations are equivalent. While our formulations (13) and (14) have the following two advantages:

1. Our formulations have fewer variables and constraints. In contrast, the eigenvalue based formulation involves 39 variables and 536 constraints. As shown in the following sections, the simplicity of our formulations will result in much more efficient solvers and enable the proof of tightness and local stability.

2. Our formulations are easy to integrate priors for each 2D-2D correspondence by simply introducing weights in the objective. For example, we may introduce weights for samples by slightly changing the objective as \( \sum_{i=1}^{N} w_i (f_i E f_i')^2 \), where \( w_i \geq 0 \) is the weight for \( i \)-th observation. In this paper, we simply set \( w_i = 1 \) without loss of generality. For general cases in which \( w_i \geq 0 \), we can keep current formulation by simply conducting variable substitutions \( f_i \leftarrow \sqrt{w_i} f_i \) and \( f_i' \leftarrow \sqrt{w_i} f_i' \).

### 4 Semidefinite Relaxation and Optimization

QCQP is a well-studied problem in the global optimization literature with many applications. Solving its general case is an NP-hard problem. Global optimization methods for QCQP are typically based on convex relaxations of the problem. There are two main relaxations for QCQP: semidefinite relaxation, and the reformulation-linearization technique. In this paper, we use SDR because it usually has better performance [30] and it is convenient for tightness analysis.

Let us consider a QCQP in a general form as

\[
\min_{x \in \mathbb{R}^n} \quad x^T C_0 x \\
\text{s.t.} \quad x^T A_i x = b_i, \quad i = 1, \ldots, m.
\]

Matrices \( C_0, A_1, \ldots, A_m \in \mathbb{S}^n \), where \( \mathbb{S}^n \) denotes the set of all real symmetric \( n \times n \) matrices. In our problem,

\[
x = \begin{bmatrix} e; t \end{bmatrix},
\]

is a vector stacking all entries in essential matrix \( E \) and translation vector \( t \); \( n = 12; m = 7; C_0 = \begin{bmatrix} C & 0_{9 \times 3} \\ 0_{3 \times 9} & 0_{3 \times 3} \end{bmatrix} \), \( A_1 \sim A_7 \) correspond to the canonical form \( x^T A_i x \) of Eqs. (15a) \(-\) (15g), respectively.

A crucial first step in deriving an SDR of problem (16) is to observe that

\[
x^T C_0 x = \text{trace}(x^T C_0 x) = \text{trace}(C_0 xx^T),
\]

\[
x^T A_i x = \text{trace}(x^T A_i x) = \text{trace}(A_i xx^T).
\]

In particular, both the objective and constraints in problem (16) are linear in the matrix \( xx^T \). Thus, by introducing a new variable \( X = xx^T \) and noting that \( X = xx^T \) is equivalent to \( X \) being a rank one symmetric positive semidefinite (PSD) matrix. Thus we obtain the following equivalent formulation of problem (16)

\[
\min_{X \in \mathbb{S}^n} \text{trace}(C_0 X) \quad \text{s.t.} \quad \text{trace}(A_i X) = b_i, \quad i = 1, \ldots, m, \\
X \succeq 0, \quad \text{rank}(X) = 1.
\]

Here, \( X \succeq 0 \) means that \( X \) is PSD. Solving rank constrained semidefinite programs (SDPs) is NP-hard [31]. SDR drops the rank constraint \( \text{rank}(X) = 1 \) to obtain the following relaxed version of problem (20)

\[
\min_{X \in \mathbb{S}^n} \text{trace}(C_0 X) \quad \text{s.t.} \quad \text{trace}(A_i X) = b_i, \quad i = 1, \ldots, m, \\
X \succeq 0.
\]

The problem (21) turns out to be an instance of semidefinite programming [31], which belongs to convex optimization and can be readily solved using primal-dual interior point methods [32]. Its dual problem is

\[
\max_{\lambda} \quad b^T \lambda \\
\text{s.t.} \quad Q(\lambda) = C_0 - \sum_{i=1}^{m} \lambda_i A_i \succeq 0,
\]

where \( b = [b_1, \ldots, b_m]^T, \lambda = [\lambda_1, \ldots, \lambda_n]^T \in \mathbb{R}^n \). Problem (22) is called the Lagrangian duality of problem (16), and \( Q(\lambda) \) is the Hessian of the Lagrangian. In our problem, \( b_f = 1, b_i = 0 \) if \( i \neq 7 \), and \( b^T \lambda = \lambda_7 \).

In summary, the relationships between the main formulations are demonstrated by Fig. 1.

**Lemma 1.** For QCQP problem (14), there is no duality gap between the primal SDP problem (21) and its dual problem (22).

**Proof.** Denote the optimal value for problem (21) and its dual problem (22) as \( f_{\text{primal}} \) and \( f_{\text{dual}} \). The inequality \( f_{\text{primal}} \geq f_{\text{dual}} \) follows from weak duality. Equality, and the existence of \( X^* \) and \( \lambda^* \) which attain the optimal values follow if we can show that the feasible regions of both the primal and dual problems have nonempty interiors, see Theorem 3.1 in [31] (also known as Slater’s constraint qualification [33].)

For the primal problem, Let \( E_0 \) be an arbitrary point in the essential matrix manifold \( \mathcal{M}_E: E_0 = [t_0] R_0, \) where \( \|t_0\| = 1 \). Denote \( x_0 = [\text{vec}(E_0); \text{vec}(t_0)] \). It can
be verified that \( X_0 = x_0 X_0^T \) is an interior in the feasible domain of the primal problem. For the dual problem, let \( \lambda_0 = [-1, -1, -1, 0, 0, 0, -3]^T \). Recall that \( C \geq 0 \) and it can be seen that \( Q(\lambda_0) = \begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix} + I \succ 0 \). That means \( \lambda_0 \) is an interior in the feasible domain of the dual problem.

According to Lemma 1, the Lagrangian duality is equivalent to its SDR for our formulation of \( N \)-point problem. The Lagrangian duality provides an important theoretical tool for analyzing tightness and local stability.

### 4.1 Essential Matrix and Relative Pose Recovery

Once the optimal \( X^* \) of the SDP primal problem (21) is calculated by an SDP solver, it leaves a task to recover the optimal essential matrix \( E^* \). First we extract a submatrix of \( X^* \) by

\[
X^*_e = X^*(1:9, 1:9). \tag{23}
\]

Empirically, we found that \( \text{rank}(X^*_e) = 1 \). Denote the eigenvector that corresponding to the nonzero eigenvalue as \( e^* \), then the optimal essential matrix is recovered by

\[
E^* = \text{mat}(e^*, [3, 3]), \tag{24}
\]

where \( \text{mat}(e, [r, c]) \) means reshape the vector \( e \) to an \( r \times c \) matrix by column-first order.

Once the essential matrix is obtained, we can recover the rotation matrix and translation vector by the standard method in textbooks [2]. In Section 4.2, the theoretical guarantee of such pose recovery method will be provided. In Section 5, the tightness that guarantees the global optimality will be provided.

### 4.2 Necessary and Sufficient Conditions for Global Optimality

The following Theorem 2 provides a theoretical guarantee for previously described pose recovery method. Denote \( X_e \) as the top-left \( 9 \times 9 \) submatrix formed by the entries of essential matrix \( E_3 \); denote \( X_t \) as the bottom-right \( 3 \times 3 \) submatrix formed by the entries of translation vector \( t \), i.e., \( X_e = X(1:9, 1:9) \) and \( X_t = X(10:12, 10:12) \).

**Theorem 2.** For QCQP problem (14), its SDR is tight if and only if: the optimal solution \( X^* = [X^*_e; X^*_t] \) to its primal SDP problem (21) satisfies \( \text{rank}(X^*_e) = \text{rank}(X^*_t) = 1 \).

**Proof.** First, we prove the if part. Note that \( X^*_e \) and \( X^*_t \) are real symmetric matrices because they are in the feasible region of the primal SDP. Besides it is given that \( \text{rank}(X^*_e) = \text{rank}(X^*_t) = 1 \), thus there exist two vectors \( e^* \) and \( t^* \) satisfying \( e^*(e^*)^T = X^*_e \) and \( t^*(t^*)^T = X^*_t \).

Since the constraints in problem (14) do not include any cross term between \( E \) and \( t \), the intersection part of \( E \) and \( t \) in matrix \( A_0 \) is zero in SDP problem. Substituting \( e^* \) and \( t^* \) into Eq. (19), it can be verified that \( e^* \) and \( t^* \) satisfy the constraints in primal problem (14). We can see \( X^* \) and its uniquely determined derivatives \( (e^* \text{ and } t^*) \) are feasible solutions for the SDP relaxation problem and the primal problem, respectively. Thus the relaxation is tight.

Then we prove the only if part. Since the SDP relaxation is tight, it has that \( \text{rank}(X^*) = 1 \). Then \( \text{rank}(X^*_e) \leq 1 \), \( \text{rank}(X^*_t) \leq 1 \), where

\[
\text{rank}(X^*_t) \leq 1.
\]

Since \( X^*_e \) and \( X^*_t \) can not be zero matrices (otherwise \( X^* \) is not in the feasible region), the equalities should hold.

Theorem 2 provides a necessary and sufficient global optimality condition to recover the optimal solution for the primal problem. Empirically, the optimal \( X^* \) by the SDP problem always satisfies this condition. Specifically, there are 2 nonzero diagonal blocks in it, which are \( X_e \) and \( X_t \), i.e.,

\[
X^* = \begin{bmatrix} X^*_e & \triangle \\ \triangle & X^*_t \end{bmatrix} = \begin{bmatrix} e^*(e^*)^T & \triangle \\ \triangle & t^*(t^*)^T \end{bmatrix}. \tag{25}
\]

The \( \triangle \) parts could be arbitrary that satisfy the matrix symmetry.

**Remark:** The block diagonal structure of \( X^* \) in Eq. (25) is caused by the sparsity pattern of the problem. The aggregate sparsity pattern in our SDP problem, which is the union of the individual sparsity patterns of the data matrices, \( \{C_0, A_1, \ldots, A_m\} \), includes two cliques: one includes the 1~9-th entries of \( x \) and the other includes the 10~12-th entries of \( x \), see Fig. 2(a). There is no common nodes in these two cliques, see Fig. 2(b). The chordal decomposition theory of the sparse SDPs can explain the structure of \( X^* \) well. The interested reader may refer [34], [35] for more details.

### 4.3 Time Complexity

First, we consider the time complexity with respect to the observation number \( N \). The construction of the optimization problem is linear with observation number, so the time complexity of problem construction is \( O(N) \). While the computation of \( C \) is rather efficient, see Eq. (12). The time complexity of optimization is independent of sample number \( N \), so the time complexity of the SDP optimization is \( O(1) \).

Second, we discuss the time complexity with respect to the variable number \( n \) and constraint number \( m \). Most convex optimization solvers handle SDPs using an interior-point algorithm. The SDP problem (21) can be solved with a worst case complexity of

\[
O(n m n^{1/2} \log(1/\epsilon)), \tag{26}
\]

give a solution accuracy \( \epsilon > 0 \) [32]. It can be seen that the time complexity can be largely reduced given smaller \( n \) and \( m \). Since our formulations have much fewer variables...
and constraints than that in [6], they own much lower time complexity. The above complexity does not assume sparsity or any special structures in the data matrices. In practice, our formulation is about two orders faster than that in [6].

Finally, we discuss the time complexity for pose recovery. In our method, the essential matrix is recovered by finding the eigenvector corresponding to the largest eigenvalue of a $9 \times 9$ matrix $X^+_e$. In contrast, the method in [6] needs to calculate 4 eigenvectors corresponding to the 4 largest eigenvalues of a $40 \times 40$ matrix. Thus our method has smaller time complexity for pose recovery.

5 Tightness and Local Stability of the SDP Relaxation

In this section, we prove the tightness and local stability of the SDR for our problem. Our proof is mainly based on [23], [24]. First, we prove that the SDR is tight given noise-free observations in our problem. The proof is based on Lemma 2.4 in [24], and it will be presented in Section 5.1. Then we prove that the SDR has local stability near the noise-free observations. The proof is based on Theorem 5.1 in [24], and will be presented in Section 5.2. To understand the proofs in this section, preliminary knowledge about optimization, manifold, and algebraic geometry are necessary. We recommend the readers to refer Chapter 6 of [23] or [24] for more details.

In the following text, the bar on a symbol stands for the upper limit of the parameter of a vector after the normalization. For example, $\bar{\mathbf{x}} \in \mathbb{R}^n$ is the vector $\mathbf{x}$ normalized to unit length.

In the following text, the bar on a symbol stands for a value under the noise-free case. For example, $\bar{\mathbf{x}} \in \mathbb{R}^n$ is the vector $\mathbf{x}$ normalized to unit length.

5.1 Tightness of the SDP Relaxation

Lemma 2. If the point observations are noise-free, the matrix $\bar{\mathbf{X}}$ in Eq. (12) satisfies that $\text{rank}(\mathbf{C}) \leq \min(N, 8)$, where $N$ is the number of point matches. The equality holds except for degenerate configurations including points on a ruled quadratic, points on aulings, and constraints than that in [6], they own much lower time complexity. The above complexity does not assume sparsity or any special structures in the data matrices. In practice, our formulation is about two orders faster than that in [6].

Finally, we discuss the time complexity for pose recovery. In our method, the essential matrix is recovered by finding the eigenvector corresponding to the largest eigenvalue of a $9 \times 9$ matrix $X^+_e$. In contrast, the method in [6] needs to calculate 4 eigenvectors corresponding to the 4 largest eigenvalues of a $40 \times 40$ matrix. Thus our method has smaller time complexity for pose recovery.

5 Tightness and Local Stability of the SDP Relaxation

In this section, we prove the tightness and local stability of the SDR for our problem. Our proof is mainly based on [23], [24]. First, we prove that the SDR is tight given noise-free observations in our problem. The proof is based on Lemma 2.4 in [24], and it will be presented in Section 5.1. Then we prove that the SDR has local stability near the noise-free observations. The proof is based on Theorem 5.1 in [24], and will be presented in Section 5.2. To understand the proofs in this section, preliminary knowledge about optimization, manifold, and algebraic geometry are necessary. We recommend the readers to refer Chapter 6 of [23] or [24] for more details.

In the following text, the bar on a symbol stands for a value under the noise-free case. For example, $\bar{\mathbf{C}}$ stands for the matrix in objective constructed by noise-free observations; $\bar{\mathbf{x}}$ is the optimal state estimated by noise-free observations; $\mathbf{t} = [\bar{t}_1, \bar{t}_2, \bar{t}_3]^\top$ is the optimal translation estimated by noise-free observations.

Lemma 3. Let $\mathbf{C} \in \mathbb{R}^{n \times n}$ be positive semidefinite. If $\mathbf{x}^\top \mathbf{C} \mathbf{x} = 0$ for a given vector $\mathbf{x}$, then $\mathbf{C} \mathbf{x} = 0$.

Proof. Since $\mathbf{C}$ is positive semidefinite, its eigenvalues are non-negative. Suppose the rank of $\mathbf{C}$ is $r$. The eigenvalues can be listed as $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 = \sigma_{r+1} = \cdots = \sigma_n$. Denote the $i$-th eigenvector as $\mathbf{x}_i$. The eigenvectors are orthogonal to each other, i.e., $\mathbf{x}_i^\top \mathbf{x}_j = 0$ when $i \neq j$. Vector $\mathbf{x}$ can be expressed as $\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{x}_i$. Thus, $\mathbf{C} \mathbf{x} = \mathbf{C} \sum_{i=1}^n \alpha_i \mathbf{x}_i = \sum_{i=1}^n \alpha_i \mathbf{x}_i^\top \mathbf{C} \mathbf{x}_i$ and $\mathbf{x}^\top \mathbf{C} \mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{x}_i^\top \mathbf{C} \mathbf{x}_i$. Given $\mathbf{x}^\top \mathbf{C} \mathbf{x} = 0$, we have $\alpha_i = 0$ for $i = 1, \ldots, r$. So $\mathbf{C} \mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{x}_i = 0$ is attained.

Lemma 4. If the point observations $\{(\mathbf{f}_i, t_i')\}_{i=1}^n$ are noise-free, there is zero-duality-gap between problem (14) and its Lagrangian duality (22).

Proof. Our proof is based on the Lemma 2.4 in [24]. Let $\bar{\mathbf{x}} = [\bar{\mathbf{e}}, \bar{\mathbf{t}}]$ in the primal problem, where $\bar{\mathbf{e}}$ and $\bar{\mathbf{t}}$ are ground truth. And let $\bar{\lambda} = 0$ in Lagrangian duality (22). The three conditions needed in Lemma 2.4 in [24] are satisfied: (i) Primal feasibility. Substituting $\bar{\mathbf{x}}$ in the primal problem, the constraints are satisfied since $\bar{\mathbf{x}}$ is ground truth and point observations are noise-free. (ii) Dual feasibility.

$$Q(\lambda) = \mathbf{C}_0 - \sum_{i=1}^n \lambda_i \mathbf{A}_i = \begin{bmatrix} \mathbf{C}_0 & \mathbf{0}_{3 \times 9} \\ \mathbf{0}_{3 \times 9} & \mathbf{0}_{3 \times 3} \end{bmatrix} \geq 0. (iii)$$

Lagrangian multiplier: $\mathbf{C}_0 - \sum_{i=1}^n \lambda_i \mathbf{A}_i \bar{\mathbf{x}} = \begin{bmatrix} \mathbf{C}_0 & \mathbf{0}_{3 \times 9} \end{bmatrix}$. Since $\bar{\mathbf{e}}$ is the ground truth, $\bar{\mathbf{e}}^\top \mathbf{C}_0 \bar{\mathbf{e}} = 0$ according Eq. (11). Recall that $\mathbf{C}_0$ is a Gram matrix and thus it is positive definite. According to Lemma 3, $\mathbf{C}_0 = 0$ is attained.

5.2 Local Stability of the SDP Relaxation

In this subsection, we will prove that our QCQP formulation has a zero-duality-gap regime when its problem parameters are perturbed (e.g., with noise in the case of sensor measurements). Following [23], [24], we will use the following notations in the remainder of this section to make notation simplicity.

- $\bar{\theta} \in \Theta$ is a zero-duality-gap parameter. In our problem, $\bar{\theta} = \{\bar{\mathbf{C}}\}$.
- Given noise-free observations, let $\bar{\mathbf{x}} \in \mathbb{R}^n$ be optimal for the primal problem, and $\bar{\lambda} \in \mathbb{R}^m$ be optimal for the dual problem. In our problem, $n = 12$ and $m = 7$. According to the proof procedure of Lemma 2, $\bar{\mathbf{x}} = [\bar{\mathbf{e}}; \bar{\mathbf{t}}]$, $\bar{\lambda} = 0$.
- Denote $\bar{Q} = Q_{\bar{\theta}}(\bar{\lambda}) \in \mathbb{S}^n$ as the Hessian of the Lagrangian at $\bar{\theta}$. In our problem, $\bar{Q} = \begin{bmatrix} \mathbf{C}_0 & \mathbf{0}_{3 \times 9} \\ \mathbf{0}_{3 \times 9} & \mathbf{0}_{3 \times 3} \end{bmatrix} \geq 0$.
- Denote $\mathbf{X}_{\bar{\theta}} = \{\mathbf{x} \in \mathbb{R}^n | h_i(\mathbf{x}) = 0, i = 1, \ldots, m\}$ as the primal feasible set given $\bar{\theta}$, and denote $\bar{\mathbf{X}} = \mathbf{X}_{\bar{\theta}}$.
- Denote $\bar{\mathbf{h}}(\mathbf{x}) = [h_1(\mathbf{x}), \ldots, h_m(\mathbf{x})]$.

In our QCQP, the objective $\bar{\mathbf{e}}^\top \mathbf{C}_0 \bar{\mathbf{e}}$ is convex with $\bar{\theta}$. However, the presence of the auxiliary variables $\bar{\mathbf{t}}$ makes the objective is not strictly convex. The Theorem 5.1 in [24] provides a framework to prove the local stability for such kinds of problems.

Theorem 3 (Theorem 5.1 in [24]). Assume that the following 4 conditions are satisfied:

- $\bar{\mathbf{h}}(\mathbf{x})$ is a smooth function.
- $\mathbf{X}_{\bar{\theta}}$ is a compact set.
- $\bar{Q}$ is positive definite.
- $\bar{Q}$ is continuous and has continuous partial derivatives.
RS (restricted Slater): There exists $\mu \in \mathbb{R}^m$ such that
$$\mu^\top \nabla h_\theta(x) = 0 \quad \text{and} \quad \left(\sum_{i=1}^m \mu_i A_{i\bar{\theta}}\right)_V > 0,$$
where $V \triangleq \{v : v \in \mathbb{R}^n | Qv = 0, \dot{x}v = 0\}$.

R1 (constraint qualification): Abadie constraint qualification $\text{ACQ}_\mathbf{x}(\bar{x})$ holds.

R2 (smoothness): $W = \{ (\theta, x) | h_\theta(x) = 0 \}$ is a smooth manifold nearby $\bar{w} \triangleq (\bar{\theta}, \bar{x})$, and $\dim_w = \dim \Theta + \dim_x \mathbf{X}$.

R3 (not a branch point): $\bar{x}$ is not a branch point of $\mathbf{X}$ with respect to $v \mapsto Qv$.

Then the SDR relaxation is tight when $\theta$ is close enough to $\bar{\theta}$. Moreover, the QCQP has a unique optimal solution $x_\theta$, and the SDR problem has a unique optimal solution $x_\theta \bar{x}_\bar{\theta}$.

Among the four conditions in Theorem 3 the RS (restricted Slater) is the main assumption, which is related to the convexity of the Lagrangian function. It corresponds to the strict feasibility of an SDP. R1~R3 are regularity assumptions which are related to the continuity of the Lagrange multipliers. In the following text, we prove that the restricted Slater and R3 are satisfied in problem (14), which build an important foundation to prove the local stability. In our problem, $A_i$ and $h_i$ is independent of $\theta$, thus $A_{i\bar{\theta}} = A_i$ and $h_{i\bar{\theta}} = h_i$.

**Lemma 5** (Restricted Slater). For QCQP problem (14), suppose $\text{rank}(C) = 8$. Then there exists $\mu \in \mathbb{R}^m$ such that
$$\mu^\top \nabla h_\theta(x) = 0 \quad \text{and} \quad \left(\sum_{i=1}^m \mu_i A_{i\bar{\theta}}\right)_V > 0,$$ where $V \triangleq \{v : v \in \mathbb{R}^n | Qv = 0, \dot{x}v = 0\}$.

**Proof.** Considering the constraints Eqs. (15a) ~ (15g) in QCQP, the gradient is
$$\nabla h_\theta(x) = \begin{bmatrix} \nabla_{v_1} h_1 & \nabla_{v_2} h_1 & \nabla_{v_3} h_1 & \nabla_t h_1 \\ \vdots & \vdots & \vdots & \vdots \\ \nabla_{v_1} m & \nabla_{v_2} m & \nabla_{v_3} m & \nabla_t m \end{bmatrix}_{\bar{\theta}}$$

$$= \begin{bmatrix} 2e_1 & 0 & 0 & -2t_2 & -2t_3 \\ 0 & 2e_2 & 0 & -2t_1 & -2t_3 \\ 0 & 0 & 2e_3 & -2t_1 & -2t_2 \\ e_2 & e_1 & 0 & t_2 & t_1 \\ e_3 & 0 & e_1 & t_3 & 0 \\ 0 & e_3 & e_2 & 0 & t_2 \\ 0 & 0 & 2t_1 & 2t_2 & 2t_3 \end{bmatrix}.$$ (27)

Let
$$\mu = -\begin{bmatrix} 1 & 1 & 1 \\ 1/2 & 1/2 & 1/2 \\ 2 & 2 & 2 \\ t_1 & t_2 & t_3 \end{bmatrix}.$$ (28)

Note that for noise-free cases
$$\bar{t}^\top \mathbf{E} = \bar{e}^\top \left(\begin{bmatrix} \bar{t} \\ \mathbf{R} \end{bmatrix}\right) = 0,$$
so it satisfies that
$$\bar{t}_1 \bar{e}_1 + \bar{t}_2 \bar{e}_2 + \bar{t}_3 \bar{e}_3 = 0.$$

Combining with this result, it can be verified that
$$\mu^\top \nabla h_\theta(x) = 0.$$

It remains to check the positivity condition. According to the definition of $V$, $\begin{bmatrix} Q \end{bmatrix}_x v = 0 \iff \begin{bmatrix} C & 0 \\ \bar{e} & \bar{t} \end{bmatrix} v = 0$. Since $C$ is constructed by noise-free observations, $\bar{C} \bar{e} = 0$. In other words, $\bar{e}$ is orthogonal to the space spanned by $\bar{C}$. It is given that $\text{rank}(C) = 8$, thus $\text{rank} \left(\begin{bmatrix} C \\ \bar{e} \end{bmatrix}\right) = 9$. Considering $v$ as a non-trivial solution of a homogeneous linear equation system, $v$ can be expressed by a coordinate system $v = \begin{bmatrix} 0_{9 \times 1} \\ t \end{bmatrix}$.

It can be seen that only $10 \sim 12$-th entries in coordinate system $v$, which correspond to $t$, are nonzero. Take Hessian for variable $t$ and calculate the linear combination with coefficient $\mu$, we have
$$\begin{align*}
A(\mu) & = \sum_{i=1}^m \mu_i \nabla^2_t h_{i\bar{\theta}}(x) \\
& = \begin{bmatrix} t_2^2 + t_3^2 - t_1 t_2 & -t_1 t_3 & -t_1 t_2 \\ -t_1 t_2 & t_1^2 + t_3^2 & -t_1 t_3 \\ -t_1 t_2 & -t_1 t_3 & t_1^2 + t_3^2 \end{bmatrix} \\
& = t_1 \bar{t}_x \begin{bmatrix} t \end{bmatrix}_x = \mathbf{E} \mathbf{E}^\top.
\end{align*}$$ (29)

Recall that in the proof procedure of Theorem 3 we have proved that the eigenvalues of $\mathbf{E} \mathbf{E}^\top$ are $1, 1, 1, 0$, and $0$. The eigenvalues of $A(\mu)$ are $1, 1, 1, 0$. And it can be verified that $t = [t_1, t_2, t_3]^\top$ is the normalized eigenvector corresponding to eigenvalue $0$ of $A(\mu)$. Thus, $V$ is the orthogonal complement of $t$.

Since for any vector $v \in V \setminus \{0\}$, its $10 \sim 12$-th entries are orthogonal to $t$, $v_{10:12}A(\mu)v_{10:12}$ is strictly positive. It follows that $\left(\sum_{i=1}^m \mu_i A_{i\bar{\theta}}\right)_V = A(\mu)|_{t} > 0$.

**Definition 1** (Branch Point [24]). Let $\pi : \mathbb{R}^n \rightarrow \mathbb{R}^k$ be a linear map: $v \mapsto Qv$. Let $X \subseteq \mathbb{R}^n$ be the zero set of the equation system $h(x) = [h_1(x), \cdots, h_m(x)]$, and let $T_x X \triangleq \ker(\nabla h(x))$ denote the tangent space of $X$ at $x$. We say that $x$ is a branch point of $X$ with respect to $\pi$ if there is a nonzero vector $v \in T_x X$ with $\pi(v) = 0$.

**Lemma 6.** For QCQP problem (14), suppose $\text{rank}(C) = 8$. Then $x = [\bar{x}; t]$ is not a branch point of $X$ with respect to $\pi : v \mapsto Qv$.

**Proof.** It can be verified that $\pi(v) = 0 \iff \bar{Q}v = 0 \iff \begin{bmatrix} C & 0_{9 \times 3} \\ 0_{3 \times 9} & 0_{3 \times 3} \end{bmatrix} v = 0 \iff v = \begin{bmatrix} \bar{c} \\ t \end{bmatrix}$, where $c$ and $t$ are free parameters. The last equivalence takes advantage of $\text{rank}(C) = 8$. If $x$ is a branch point, there should exist a nonzero vector $v \in \ker(\nabla h(x))$, i.e., $\nabla h(x)v = 0$. We will prove that such nonzero $v$ does not exist. Substituting $v = \begin{bmatrix} \bar{c} \\ t \end{bmatrix}$ into equation $\nabla h_\theta(x)v = 0$ (see Eq. (27)), we obtain a homogeneous linear system with unknowns $t = [t_1, t_2, t_3]^\top$ and
$$\begin{align*}
\bar{e}_1 \bar{t}_1 c - \bar{t}_2 t_3 - \bar{t}_3 t_2 &= 0, \\
\bar{e}_2 \bar{t}_1 c - \bar{t}_1 t_2 - \bar{t}_3 t_3 &= 0, \\
\bar{e}_3 \bar{t}_1 c - \bar{t}_1 t_2 - \bar{t}_2 t_3 &= 0, \\
2\bar{e}_1 \bar{t}_2 c + \bar{t}_1 t_2 + t_1 t_3 &= 0, \\
2\bar{e}_2 \bar{t}_3 c + \bar{t}_2 t_3 + t_1 t_3 &= 0, \\
\bar{t}_1 t_1 + \bar{t}_2 t_2 + \bar{t}_3 t_3 &= 0.
\end{align*}$$ (30)

Eliminating $t$ from Eqs. (30a) (30b) (30c) (30d), we obtain that $(\bar{e}_1 \bar{t}_1 c + \bar{e}_2 \bar{t}_2 c + \bar{e}_3 \bar{t}_3 c) = 0$ and $c = 0$. Substituting $c = 0$ into this equation system, it can be further verified that this equation system only has zeros as its solution. So $v$ can only be a zero vector. □
Theorem 4. For problem (14) and its Lagrangian duality (22), let C being constructed by noise-free observations and rank(C) = 8. (i) There is zero-duality-gap whenever C is close enough to C. In other words, there exists a hypersphere of nonzero radius ε with center C, i.e., B(C, ε) = ∥C − C∥ ≤ ε, such that for any C ∈ B(C, ε) there is zero-duality-gap. (ii) Moreover, the SDP relaxation problem recovers the optimum of original problem.

Proof. From Lemma 4 when the point observations are noise-free, the relaxation is tight. From the proof procedure, the optimum is x = [e; t], which uniquely determines the zero-duality-gap parameter C. Our proof is an application of the Theorem 3. The four conditions needed in Theorem 3 are satisfied: (R5) From Lemma 5, the restricted Slater is satisfied. (R1) The equality constraints in the primal problem forms a variety. Abadie constraint qualification (ACQ) holds everywhere since the variety is smooth and the ideal is radical, ref. Lemma 6.1 in [24]. (R2) Smoothness. The normalized essential matrix manifold is smooth [26]. (R3) From Lemma 6 x is not a branch point of X.

Now we complete the proof that the SDR of our proposed QCQP problem is tight under low noise observations. In other words, when the observation noise is small, Theorem 3 guarantees that the optimum of the original QCQP can be found by optimizing its SDR. Our proof is an application of the method in [23], [24], while the proof is non-trivial. First, we proposed a compact parameterization without any redundant constraint which enables the proof. Second, we prove the conditions are satisfied by construction. To the best of our knowledge, this is the first time to achieve such kinds of theoretical results for the SDR in N-point problem. Also, it is the first application of Theorem 3 for practical problems. It should note that finding the noise bounds that SDR can tolerate is still an open problem. In practice, we found that the SDR is tight even for large noise level which is much larger than that in actual occurrence.

6 Experimental Results

We compared our method against several state-of-the-art methods on synthetic and real data. Specifically, we compared our method to 6 classical or state-of-the-art methods:

- general five-point method 5pt-Nister [1] for relative pose estimation.
- two general methods for fundamental matrix estimation, including seven-point method 7pt [2] and eight-point method 8pt [37].
- state-of-the-art eigensolver method eigensolver proposed by Kneip and Lynen [5] and a certifiably globally optimal solution by Briales et al. [6] which is referred to as SDP-Briales. These two methods and ours have exactly the same optimization objective.
- two view bundle adjustment (BA) method nonlinear.

Among these methods, the implementation of SDP-Briales was provided by the authors. And the implementations of other comparison methods were provided by OpenGV [58]. The method eigensolver and nonlinear need initialization, which we initialized using 8pt for all point correspondences.

To evaluate the performance of the proposed algorithms we separately compare the relative rotation and translation accuracy. Specifically, we define

- the angle difference for rotations as

\[ \varepsilon_{\text{rot}}(\text{degrees}) = \arccos \left( \frac{\text{trace}(\hat{R}_\text{true} R^*) - 1}{2} \right) \cdot \frac{180}{\pi}, \]

- and the translation direction error as

\[ \varepsilon_{\text{tran}}(\text{degrees}) = \arccos \left( \frac{t_\text{true}^\top t^*}{\|t_\text{true}\| \cdot \|t^*\|} \right) \cdot \frac{180}{\pi}. \]

In the above criteria, R_true and R* are ground truth and estimated rotation, respectively; t_true and t* are ground truth and estimated translation, respectively.

6.1 Efficiency and SDP Solver Selection

We test two free interior point method (IPM) solvers, including SDPT3 [39] and SeDuMi [40]. Default parameters are used for all solvers in the following experiments. Our method is implemented in MATLAB, and all experiments are performed on an Intel Core i7 CPU with 2.40 GHz.

The two IPM solvers generate similar results, while SDPT3 has slightly better accuracy than SeDuMi. The SDP optimization takes about 48 ms with SeDuMi and 110 ms with SDPT3. Specifically, the problem construction with 100 observations takes about 3 ms; the essential matrix recovery takes about 3 ms; pose decomposition take about 8 ms. Totally it takes 62 ms to solve an N-point problem with SeDuMi. Thus our method is much more efficient than SDP-Briales, which takes about 5.19 seconds in total. The comparison of ours and SDP-Briales is listed in Table 1.

| method      | #variable | #constraint | total time |
|-------------|-----------|-------------|------------|
| SDP-Briales | 40        | 536         | 5.19 s     |
| ours        | 12        | 7           | 62 ms      |

Since there are some specific structures in our problem, such as matrices \(\{A_i\}_{i=1}^m\) are sparse, and rank(\(X^*\)) is known a priori to be small relative to the dimension \(n\) of the problem. There are some SDP solvers which exploit the structures of the SDPs, such as sparsity of the data and the low-rank property of the solution, to accelerate the optimization [41], [42], [43]. Their efficiency has been successfully validated for large-scale problems. While the SDP in our method is small-scale, and these solvers are not necessarily more efficient than general solvers. We tested the SparseCoLO [41], MoDuMi (Modified SeDuMi for large-and-sparse low-rank SDPs) [42], and CDCS (Conic Decomposition Conic Solver) [43].

Compared with SeDuMi, we found that: (i) SparseCoLO takes about 64 ms, which is slightly slower. (ii) MoDuMi takes 1.7 seconds, which is much slower. (iii) CDCS is faster than SeDuMi (25 ms), but its accuracy is unsatisfactory. Based on these results, we use SeDuMi in the following experiments. Pursuing a customized more efficient solver is a promising direction for future work.
nonlinear local gradient descent of ours performs slightly better than the mean of $\varepsilon^2$ functions for executions time are demonstrated in Fig. 3. The results for all methods with varying image noise levels are shown in Fig. 4. We have the following observations. (1) When the noise level is below 2 pixels, eigensolver, nonlinear and ours have similar performance, and they obviously outperform other methods. (2) When the noise level is over 2 pixels, from the mean of $\varepsilon_{\text{ran}}$ and $\varepsilon_{\text{ran}}$, ours has less oscillation than other methods. This is due to the property that our method does not depend on any initialization and can find the global optima. In contrast other two methods need a good initial value. (3) When the noise level is over 2 pixels, from the median of $\varepsilon_{\text{ran}}$ and $\varepsilon_{\text{ran}}$, eigensolver and nonlinear performs slightly better than ours. This means that the local gradient descent of eigensolver has better accuracy given good initialization. The nonlinear which minimizes the geometric error achieves the best performance given good initialization. In this experiment, the performance gap between geometric error based method (nonlinear) and two best algebraic error based methods (eigensolver and ours) is small.

Next, we fix the image noise level to 0.5 pixel in standard deviation and vary the number $N$ of point correspondences. The methods 5pt-Nister, 7pt, and 8pt can only take a subset of the point correspondences. The methods eigensolver, nonlinear, and ours take advantage of all the point correspondences. To make a fair comparison, first we randomly sample required number of point correspondences for 5pt-Nister, 7pt, and 8pt, and repeat 20 times for each method. Then we find the optimal relative pose among them. Since all point correspondences are inliers, we can not use the inlier number as the criterion to find the optimal rotation as traditional RANSAC framework. Instead, we use the algebraic error to find the optimal relative rotation for each method. Once the relative rotation has been obtained, we can calculate the translation $t$. Recall that $(f, f')$ represent point correspondences across two images. The epipolar constraint can be written as

$$f_i^T[t]xRf'_i = 0.$$  

(31)

Since $R$ has been calculated already, each point correspondence provides a linear constraint on the entries of the translation vector $t$. Since the relative translation can be only determined up to an unknown scale factor, and therefore has only two DoFs. After a direct linear transformation, a normalized version of $t$ can be recovered by simple linear derivation of the right-hand null-space vector (e.g. via singular value decomposition). Given $N$ ($N \geq 2$) point matches, the least squares fit of $t$ can be determined by considering the singular vector corresponding to the smallest singular value.

The pose estimation results are shown in Fig. 5. We have the following observations. (1) The errors of all methods decrease when increasing the numbers of point correspondences. (2) The eigensolver, nonlinear, and ours have similar performance. When $N > 20$, these three methods are consistently better than other methods. (3) The nonlinear have the best performance regardless of the viewpoint. There is a belief that geometric error have better performance than algebraic error [2]. From our experiments, confirmed this conclusion, and we found that the results of these two error metrics does not have too much difference when $N$ is large.

### 6.3 Results with Real Data

In order to demonstrate the applicability of our relative pose estimation in real-world scenarios, we present a frame-to-frame visual odometry (VO) pipeline relying on the proposed $N$-point method. We compare our method against other alternatives by embedding different pose estimator into the pipeline. The comparison methods includes 5pt-Nister [1], eigensolver [3], and nonlinear bundle adjustment. The visual odometry pipeline essentially follows the paradigm introduced in [14]. The brief outline of the steps executed for each frame is given as follows:

- Extract point features using ORB [15]. Match the features to a previous frame while ensuring sufficiently large
disparity (keyframes are defined whenever an average disparity measure exceeds a certain threshold).
- Estimate the relative pose using preemptive RANSAC [3]. Hypotheses are generated by the 5pt-Nister. Optionally, eigensolver, nonlinear, and ours take advantage of the inliers preserved by 5pt-Nister to re-estimate the relative pose.
- Estimate the relative scale for the current pose estimate and attach it to the last estimate.

We evaluate the performance on the EuRoC MAV dataset [46]. This dataset was collected by an MAV operating in indoor environments. It contains stereo images from a global shutter camera running at 20 FPS and synchronized IMU measurements captured at 200 Hz. Each dataset comes with ground truth captured by an external VICON motion capture system. Sample images from this dataset are shown in Fig 6.

For quantitative analysis, we evaluate the pose accuracy between successive keyframes by $\varepsilon_{\text{rot}}$ and $\varepsilon_{\text{tran}}$. Table 2 demonstrates the median and mean of $\varepsilon_{\text{rot}}$ and $\varepsilon_{\text{tran}}$ for translation and rotation, respectively. From this experiment, we found that ours and eigensolver have consistently better performance than classical 5pt method. The method nonlinear generates better results than its initialization for most of the time, while it occasionally converges to much worse results. The eigensolver has the best performance if its initialization is good and it converges to the global optima. In contrast, our method does not need any initialization and can always find a global optimum. Since it uses all the inliers successfully, this leads to noise cancellation and highly accurate relative pose estimation.

7 Conclusions
The present paper introduces a new way to formulate and solve the non-minimal essential matrix estimation. First, we reformulate this problem as a simple QCQP by providing an equivalent form of the essential matrix manifold. Second, a semidefinite relaxation is adopted to convert this problem to an SDP problem, which belongs to convex optimization. Third, pose recovery from an optimal solution for SDP is proposed. Finally, theoretical analysis of tightness and local stability are provided. Our method is stable, provably optimal, and relatively easy to implement. Experiments demonstrate that our approach always finds and certifies (a-posteriori) the global optimum of the optimization problem,
and it is dozens of times faster than state-of-the-art globally optimal solutions.

**APPENDIX A**

**ANOTHER FORMULATION OF N-POINT PROBLEM**

The following Theorem also provides an equivalent condition to define $\mathcal{M}_E$, which will simplify the optimization in our method.

**Lemma 7.** For an essential matrix $E$ which can be decomposed by $E = [t] \times R$, it satisfies that $\text{trace}(E^T E) = \sum_{i=1}^{3} \sum_{j=1}^{3} E_{ij}^2 = 2\|t\|^2$.

Proof. Note that the norm of each row of $R$ is 1, and the rows of $R$ are orthogonal to each other. Taking advantage of $E = [t] \times R$, it can be verified that $\text{trace}(E^T E) = \sum_{i=1}^{3} \sum_{j=1}^{3} E_{ij}^2 = 2(t_1^2 + t_2^2 + t_3^2) = 2\|t\|^2$.

**Theorem 5 (Proposition 7.3 in [28]).** A real $3 \times 3$ matrix, $E$, is an essential matrix if and only if it satisfies the equation:

$$
EE^T - \frac{1}{2} \text{trace}(E^T E)E = 0. 
$$

**Theorem 6.** A real $3 \times 3$ matrix, $E$, is an essential matrix in $\mathcal{M}_E$ if and only if it satisfies the following two conditions:

$$
\begin{align*}
(i) \quad & \text{trace}(E^T E) = 2 \\
(ii) \quad & EE^T E = E.
\end{align*}
$$

Proof. For the $if$ direction, by combining conditions (i) and (ii) we obtain Eq. (32). According to Theorem 5, $E$ is a valid essential matrix. So there exist (at least) a pair of $t \in \mathbb{R}^3$ and $R \in \text{SO}(3)$ such that $E = [t] \times R$. According to condition (i) and Lemma 7, we have $\text{trace}(E^T E) = 2\|t\|^2 = 2$, which means $\|t\| = 1$. So we obtain $E \in \mathcal{M}_E$.

For only if direction, since $E \in \mathcal{M}_E$, it is straightforward that condition (i) is satisfied according to Lemma 7. Besides, Eq. (32) is satisfied since $E$ is an essential matrix. By substituting condition (i) in Eq. (32), we obtain condition (ii).

According to Theorem 6, an equivalent form of minimizing the algebraic error is

$$
\min_{E} \ e^T Ce 
$$

s.t. $\text{trace}(E^T E) = 2$, $EE^T E - E = 0$.

By introducing an auxiliary matrix $G$, this problem can be reformulated as a QCQP problem

$$
\min_{E, G} \ e^T Ce 
$$

s.t. $G = EE^T$, $\text{trace}(G) = 2$, $GE - E = 0$.

Note that $G$ is a symmetric matrix which introduces 6 unknowns. Thus there are 15 variables and 16 constraints in this QCQP problem.

**ACKNOWLEDGMENTS**

The author would like to thank Prof. Laurent Kneip at ShanghaiTech and and Prof. Qian Zhao at Xi’an Jiaotong University for fruitful discussions. The author also thanks Dr. Jesus Briales at the University of Malaga for providing the code of [6] and Dr. YiJie He for his help in the experiments.

**REFERENCES**

[1] D. Nister, “An efficient solution to the five-point relative pose problem,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 6, pp. 756–770, 2004.

[2] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cambridge University Press, 2003.

[3] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography,” Communications of the ACM, vol. 24, no. 6, pp. 381–395, 1981.

[4] G. Chesi, “Camera displacement via constrained minimization of the algebraic error,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 2, pp. 370–375, 2009.

[5] L. Kneip and S. Lyven, “Direct optimization of frame-to-frame rotation,” in IEEE International Conference on Computer Vision, 2013, pp. 2352–2359.

[6] J. Briales, L. Kneip, and J. Gonzalez-Jimenez, “A certifiably globally optimal solution to the non-minimal relative pose problem,” in IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 145–154.

[7] V. Lepetit, F. Moreno-Noguer, and P. Fua, “EPnP: An accurate $O(n)$ solution to the PnP problem,” International Journal of Computer Vision, vol. 81, no. 2, pp. 155–166, 2009.

[8] K. S. Arun, T. S. Huang, and S. D. Blostein, “Least-squares fitting of two 3-D point sets,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 9, no. 5, pp. 698–700, 1987.

[9] B. K. Horn, “Closed-form solution of absolute orientation using unit quaternions,” Journal of the Optical Society of America A, vol. 4, no. 4, pp. 629–642, 1987.

[10] Q. Cai, Y. Wu, L. Zhang, and P. Zhang, “Equivalent constraints for two-view geometry: Pose solution/pure rotation identification and 3D reconstruction,” International Journal of Computer Vision, vol. 127, no. 2, pp. 163–180, 2019.

[11] R. I. Hartley and F. Kahl, “Global optimization through rotation space search,” International Journal of Computer Vision, vol. 62, no. 1, pp. 64–79, 2009.

[12] T. Migita and T. Shakanaga, “Evaluation of epipole estimation methods with/without rank-2 constraint across algebraic/geometric error functions,” in IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1–7.

[13] R. Hartley, “Minimizing algebraic error in geometric estimation problem,” in IEEE International Conference on Computer Vision, 1998, pp. 469–476.

[14] G. Chesi, A. Garulli, A. Vicino, and R. Cipolla, “Estimating the fundamental matrix via constrained least-squares: A convex approach,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 3, pp. 397–401, 2002.

[15] Y. Zheng, S. Sugimoto, and M. Okutomi, “A branch and contract algorithm for globally optimal fundamental matrix estimation,” in IEEE Conference on Computer Vision and Pattern Recognition, 2011, pp. 2953–2960.

[16] F. Bagaria, A. Bartoli, D. Henrion, J.-B. Lasserre, J.-J. Orteu, and T. Sentenac, “Rank-constrained fundamental matrix estimation by polynomial global optimization versus the eight-point algorithm,” Journal of Mathematical Imaging and Vision, vol. 53, no. 1, pp. 42–60, 2015.

[17] M. Mevissen and M. Kojima, “SDP relaxations for quadratic optimization problems derived from polynomial optimization problems,” Asia-Pacific Journal of Operational Research, vol. 27, no. 1, pp. 15–38, 2010.

[18] F. I. Kahl and D. Henrion, “ Globally optimal estimates for geometric reconstruction problems,” International Journal of Computer Vision, vol. 74, no. 1, pp. 3–15, 2007.

[19] C. Olsson and A. Eriksson, “Solving quadratically constrained geometrical problems using Lagrangian duality,” in International Conference on Pattern Recognition. IEEE, 2008, pp. 1–4.

[20] C. Ahlert, S. Agarwal, and R. Thomas, “A QCQP approach to triangulation,” in European Conference on Computer Vision. Springer, 2012, pp. 654–667.

[21] L. Carlone, D. M. Rosen, G. Calafiore, J. J. Leonard, and F. Dellaert, “Lagrangian duality in 3D SLAM: Verification techniques and optimal solutions,” in IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2015, pp. 125–132.

[22] D. M. Rosen, L. Carlone, A. S. Bandeira, and J. J. Leonard, “SE-Sync: A certifiably correct algorithm for synchronization over the special Euclidean group,” International Journal of Robotics Research, 2019.
D. Cifuentes, “Polynomial systems: Graphical structure, geometry, and applications,” Ph.D. dissertation, Massachusetts Institute of Technology, 2018.

D. Cifuentes, S. Agarwal, P. A. Parrilo, and R. R. Thomas, “On the local stability of semidefinite relaxations,” arXiv preprint arXiv:1710.04287v2, 2018.

U. Helmke, K. Hupé, P. Y. Lee, and J. Moore, “Essential matrix estimation using Gauss-Newton iterations on a manifold,” International Journal of Computer Vision, vol. 74, no. 2, pp. 117–136, 2007.

R. Tron and K. Daniilidis, “The space of essential matrices as a Riemannian quotient manifold,” SIAM Journal on Imaging Sciences, vol. 10, no. 3, pp. 1416–1445, 2017.

O. D. Faugeras and S. Maybank, “Motion from point matches: multiplicity of solutions,” International Journal of Computer Vision, vol. 4, no. 3, pp. 225–246, 1990.

O. Faugeras, Three-dimensional computer vision: a geometric viewpoint. MIT press, 1993.

K. M. Anstreicher, “Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming,” Journal of Global Optimization, vol. 43, no. 2-3, pp. 471–484, 2009.

L. Vandenberghe and S. Boyd, “Semidefinite programming,” SIAM Review, vol. 38, no. 1, pp. 49–95, 1996.

Y. Ye, Interior Point Algorithms: Theory and Analysis. Wiley & Sons, 1997.

S. Boyd and L. Vandenberghe, Convex Optimization. Cambridge University Press, 2004.

M. Fukuda, M. Kojima, K. Murota, and K. Nakata, “Exploiting sparsity in semidefinite programming via matrix completion I: General framework,” SIAM Journal on Optimization, vol. 11, no. 3, pp. 647–674, 2001.

L. Vandenberghe and M. S. Andersen, “Chordal graphs and semidefinite optimization,” Foundations and Trends in Optimization, vol. 1, no. 4, pp. 241–433, 2015.

S. Maybank, “The projective geometry of ambiguous surfaces,” Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences, vol. 332, no. 1623, pp. 1–47, 1990.

R. I. Hartley, “In defence of the 8-point algorithm,” in International Conference on Computer Vision, 1995, pp. 1064–1070.

L. Kneip and P. Furgale, “OpenGV: A unified and generalized approach to real-time calibrated geometric vision,” in IEEE International Conference on Robotics and Automation, 2014, pp. 1–8.

K.-C. Toh, M. J. Todd, and R. H. Tütüncü, “SDPT3 – a MATLAB software package for semidefinite programming, version 1.3,” Optimization Methods and Software, vol. 11, no. 1-4, pp. 545–581, 1999.

J. F. Sturm, “Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones,” Optimization Methods and Software, vol. 11, no. 1-4, pp. 625–653, 1999.

S. Kim, M. Kojima, M. Mevissen, and M. Yamashita, “Exploiting sparsity in linear and nonlinear matrix inequalities via positive semidefinite matrix completion,” Mathematical programming, vol. 129, no. 1, pp. 33–68, 2011.

R. Y. Zhang and J. Lavaei, “Modified interior-point method for large-and-sparse low-rank semidefinite programs,” in IEEE 56th Annual Conference on Decision and Control. IEEE, 2017, pp. 5640–5647.

Y. Zheng, G. Fantuzzi, A. PapachristodouLou, P. Goulard, and A. Wynn, “Chordal decomposition in operator-splitting methods for sparse semidefinite programs,” Mathematical programming, Series A, 2019.

D. Nistér, O. Naroditsky, and J. Bergen, “Visual odometry,” in IEEE Conference on Computer Vision and Pattern Recognition, 2004.

E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: an efficient alternative to SIFT or SURF,” in International Conference on Computer Vision, 2011, pp. 2564–2571.

M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, “The EuRoC micro aerial vehicle datasets,” International Journal of Robotics Research, vol. 35, no. 10, pp. 1157–1163, 2016.

J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, “A benchmark for the evaluation of RGB-D SLAM systems,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, 2012, pp. 573–580.