3-D Relative Localization of Mobile Systems Using Distance-Only Measurements via Semidefinite Optimization

BOMIN JIANG
Massachusetts Institute of Technology, Cambridge, MA, USA

BRIAN D. O. ANDERSON, Life Fellow, IEEE
Hangzhou Dianzi University, Hangzhou China and The Australian National University, Canberra, Australia

HATEM HMAM
Defence Science and Technology Group, Edinburgh, Australia

In a network of cooperating unmanned aerial vehicles (UAVs), individual UAVs usually need to localize themselves in a shared and generally global frame. This paper studies the localization problem for a group of UAVs navigating in three-dimensional space with limited shared information, viz., noisy distance measurements are the only type of interagent sensing that is available, and only one UAV knows its global coordinates, the others being GPS denied. Initially, for a two-agent problem, but easily generalized to some multiagent problems, this paper first establishes constraints on the minimum number of distance measurements required to achieve the localization. This paper then proposes a composite algorithm based on semidefinite programming (SDP) in a first step, followed by maximum likelihood estimation using gradient descent on a manifold initialized by the SDP calculation. The efficacy of the algorithm is verified with experimental noisy flight data.

I. INTRODUCTION

Research on multiagent systems has been attracting increasing interest in recent years. This is due to the potential applications of this research in a variety of fields, including sensor networks [1], distributed power grids [2], distributed computation, and so on. For certain multiagent systems, such as a cooperating group of unmanned aerial vehicles (UAVs), there are some tasks, e.g., formation shape control [3], state consensus [4], or self-localization [1], which often need to be performed in a distributed way and using limited sensing.

A. Broad Problem Statement With Rationale

A common task for UAV formations is that of localizing a target, such as a radio frequency emitter. A prerequisite for this is for all agents to know their positions in a common possibly global coordinate basis, since otherwise vehicles cannot sensibly process and fuse targeting information obtained by other vehicles. However, not all vehicles may have access to GPS, because of, for example, obstacles obstructing the line of sight to satellites (e.g. receivers in a dense urban canyon) or GPS signal interference (jamming). Nevertheless, if one vehicle does have such access, then it can assist the other vehicles in the same formation to geolocate themselves, making possible collective localization of an emitter. Those other vehicles, being GPS-denied or not GPS-equipped, are assumed to be equipped with inertial navigation sensing, which means that each views its position and motion in a local coordinate frame whose orientation and position relative to the global frame are unknown. The purpose of this paper is to present an approach that expresses the local UAV position data in a global reference system, determined by the single UAV, which has access to GPS signals. Such problems are not completely novel. The book chapter [5] provides a high-level introduction to localization in a three-dimensional (3-D) environment, including limited consideration of mobility. While localization can be easily achieved in principle if vehicles have sensors delivering relative positions (i.e., both range and bearing) of their neighbors, and an ability to communicate with their neighbors, distance-only sensing is preferred in many scenarios [3], [6], [7] due to its reliability, low cost, low power, and lightweight sensors. We note, however, that bearing-only measurements are often preferred when passive sensing is required. Early analysis of bearing-only localization can be found in [8], and also, Ye et al. [9] discuss the mobile localization problem with bearing-only measurements. It is evident though that a method that can localize non-GPS-equipped UAVs using a set of distance-only measurements between mobile vehicles together with inertial navigation system (INS) data from non-GPS-equipped vehicles is of great appeal in many applications. Providing such a method is the motivation for this paper.

Additionally, we seek a method that degrades gracefully, rather than collapses, in the presence of increasing noise. Quite often in nonlinear estimation problems, estimators can perform well in small noise, but when the noise
reaches a certain threshold, collapse in performance occurs with the error of estimates sharply increasing. With the performance metric we introduce in Section IV-A, we find that no such collapse occurs with our method.

B. Relation to Other Relevant Work

In considering other relevant work, it is useful to distinguish between the application context and the methodologies. One paper with overlapping authorship, viz. [10], reports work on a rather similar problem, conducted in parallel with the work of this paper, and differing from it through the assumption of direction rather than distance measurement availability. By way of more general comment, in the context of our application, understanding the tradeoffs between numbers of measurements and signal-to-noise ratio (SNR) is critical, as is the identification of the minimum measurement requirements, since for reasons of maintaining coverty and time window of availability for measurements, often there will be a desire, or even requirement, to minimize the number of measurements. These matters are little addressed in other work. Other unique aspects of our application context in comparison with other work focused on similar applications include the treatment of fixed time synchronization offsets, consideration of adverse geometries, and the establishment of a framework for treating problems with more than two vehicles.

For the sake of completeness, we remark that there are works using very different approaches in the utilization of distance-only measurements in multiagent systems to achieve localization of vehicles whose global coordinates are not measured by the GPS; most of them require the vehicles, playing the role of mobile sensors, to move in certain standard trajectory patterns, e.g., [6] and [11]. The trajectories are adopted simply for the purpose of enabling self-localization of each agent in a global coordinate basis, or localization of a neighbor agent in a common coordinate basis; as such, they may conflict with achieving formation objectives such as surveillance of a target. Fundamentally too, this setting risks being energy inefficient. In another direction again, some work in the robotics literature, perhaps originally motivated by the Stewart–Gough platform (see [12]), studies the problem of determining the relative position and orientation between two stationary robots [13] given distance measurements between each one of a set of points on one robot with a corresponding point in a set on the other robot. The publication [13] does some pioneering work in this area and investigates this problem in 3-D, with interagent distances assumed to be measurable with high accuracy. However, this method is limited to two robots and thus when transposed to a UAV setting is only applicable to agent pairs, not to multiagent formations. It also requires a minimum of ten measurements. In fact, solving this 3-D localization problem is similar to solving the forward kinematics of a Stewart–Gough platform [14] and equivalent with merging globally rigid frameworks [15], for which seven measurements are known to be necessary and sufficient. The details of this similarity and equivalence are discussed in Section II-B. Furthermore, noise robustness is not high for the approach of [13]; one of the suggested mechanisms for coping with a large number of measurements involves solving the problem initially with a modest number of measurements before incorporating the remaining measurements. Modest noise can defeat such an approach, since the initial solution with limited measurements may be too unreliable as a starting point before integrating further measurements.

We now move on to consider aspects of our composite algorithm and its relation to the literature. Many authors have recognized that localization problems can often be formulated as nonconvex optimization problems, frequently involving a quadratic performance index and quadratic constraints, as turns out to be the case with our problem (see, e.g., [16]–[20]). Different methods have been proposed to solve them. Key steps in our algorithm include use of semidefinite programming (SDP) relaxation followed by a maximum likelihood optimization using gradient descent on the manifold of orthogonal matrices, with intermediate use of a Procrustes orthogonal matrix approximation step. Furthermore, our approach can accommodate redundant measurements, simplifying the nature of the computations thereby. On the algorithm side, there has been significant use of SDP in related problems, including sensor network localization (see, e.g., [21] and [22]). Several authors have used SDP methods in an algorithm for localization of a single point (equivalently a stationary agent) (see, e.g., the work [19]), and recently, SDP has been applied to localizing a rigid body, given (noisy) distance measurements from a collection of anchors to sensors on the body (every sensor–anchor distance being assumed known) [16]. Again, all sensors are assumed stationary.

Unlike past works and research developments, our approach to global localization is marked or guided by the following. First, it is much harder to use algorithms incorporating maximum likelihood optimization, which perform satisfactorily, as ours do, when only a few noisy measurements are available. The best possible initialization should be sought using the SDP and Procrustes steps. Furthermore, and crucially to this end, we point out the paradoxical importance of ensuring solution accuracy by including in the SDP problem formulation several additional quadratic constraints beyond the minimum; these constraints may be functionally dependent on already included constraints, but are not linearly dependent on them. Second, for the sake of completeness, we note that Procrustes algorithms (a minor component of the overall algorithm) have also been suggested by other authors for inclusion in algorithms for localization problems. Third, so far as we can establish, no papers have been able to formulate a steepest descent algorithm directly on the manifold of orthogonal matrices, such as we have, in the final stage of our algorithm or related algorithms. Finally and uniquely, our algorithms are straightforwardly expandable to include and mitigate a significant practical factor: the occurrence of a systematic but unknown bias in distance measurements arising from offsets in time synchronization.
To sum up, the key distinguishing points in our proposed method are the following.

1) Agents are not required to move in any designated pattern.
2) Only one GPS-equipped moving agent is required.
3) The only interagent measurements are distances.
4) As verified by real data, the algorithm has great robustness in noise.
5) The algorithm can handle systematic bias in the distance measurements.
6) The algorithm can function with as few as seven scalar measurements.

In short, this paper provides a practical procedure to achieve localization using distance-only measurements.

C. Organization of This Paper

Section II poses the problem of localizing in 3-D space an agent without GPS using distance measurements to a second GPS-equipped agent. It also discusses the minimal number of measurements required, including insights from the theory of graph rigidity. Section III is the most comprehensive section of this paper; it explains how the problem can be solved, using several major steps. These steps include SDP followed by a Procrustes algorithm to yield an orthogonal matrix, which gives an initial solution, and determination of a maximum likelihood estimate via a gradient algorithm initialized by the Procrustes-modified SDP solution. Issues of gradient flow on a manifold are also discussed. In addition, the existence of certain special “singular” trajectories for which a solution does not exist is discussed. This section also discusses how to deal with timing errors due to unsynchronized clocks. Simulations and results using flight test data are presented in Section IV. Concluding remarks are provided in Section VI.

II. 3-D LOCALIZATION PROBLEM AND NUMBER OF MEASUREMENTS REQUIRED

In this section, we first introduce the general localization problem for an agent pair moving arbitrarily in 3-D space. Then, we determine a minimal number of measurements required to achieve localization in the noiseless case, showing how the number is given by graph rigidity theory.

A. Problem Formulation

Consider two agents in 3-D space, as shown in Fig. 1. Agent \( x \) is the reference agent, whose position in the global coordinate system is denoted by \( p_x = [x_1, x_2, x_3]^\top \). Agent \( y \) is the agent to be localized, whose position in its own local coordinate system is denoted by \( p_y = [y_1, y_2, y_3]^\top \). Suppose further that a coordinate system transformation, including a \( 3 \times 3 \) rotation matrix \( R \) and a \( 3 \times 1 \) translation vector \( T \), can transfer agent \( y \)’s local coordinates into global coordinates, so that the position of agent \( y \) in the global coordinate system is \( Rp_y + T \). The matrix \( R \) and vector \( T \) are assumed to be constant and unknown through the interval over which measurements are taken, i.e., the cumulative drift in the INS system over the time interval of measurements is assumed negligible. Equivalently, the time interval is sufficiently short.

In addition, suppose \( d \) is the distance between the two agents and define

\[
\bar{p} = Rp_y + T - p_x.
\]

Thus, \( \bar{p} \) is the relative position vector of agent \( y \) as seen from agent \( x \), expressed in global coordinates. Immediately, we have

\[
d = \|\bar{p}\| = \|Rp_y + T - p_x\|.
\]  \hfill (1)

Suppose at time instants \( \tau = \tau_k, k = 1, 2, \ldots, N \), each agent acquires the distance measurement \( d[k] \), and each knows the coordinates \( p_x[k] \) and \( p_y[k] \) at these time instants through message interchange. The objective is to infer \( R \) and \( T \) after a finite number of measurements. Knowledge of \( R \) and \( T \) enables one to estimate the positions of agent \( y \) in global coordinates, i.e., solves the localization problem for that agent.

Note that there are six degrees of freedom to pin down, three each associated with \( R \) and \( T \). Therefore, one must expect that at least six measurements would be needed to solve the relevant equation set. With precisely six measurements, as summarized in [12], there are generally 40 solutions,\(^1\) all of which can be real and correspond to real postures. One disambiguates only by using further measurements. In the next subsection, we argue that only one further measurement is needed and, in fact, explain how the literature on graph rigidity demonstrates this point clearly.

\(^1\)The equation set is actually a set of simultaneous polynomial equations. Unless all the equations are linear, as a matter of algebra, \( n \) simultaneous polynomial equations in \( n \) unknowns always have multiple solutions, provided they are solvable in the first place.
B. Identifying the Minimal Number of Measurements Required

There are some previous papers providing some analysis of the problem of specifying the minimal number of measurements for the two agent case. They indeed solve a simplified version of this problem posed in 2-D rather than 3-D. The paper [23] points out that the two-agent problem in 2-D is almost always solvable with no less than four interagent measurements, followed by a computationally effective solution of the problem in a 2-D ambient space via the concept of a four-bar chain mechanism in mechanics.

Because the solution of this two-agent localization problem in 3-D is analogous to solving the forward kinematics of a Stewart–Gough platform [12] without the constraints that both the fixed and mobile platforms are planar, we know that with six distance measurements, the elementary two-agent localization problem stated formally in Section III normally has 40 distinct solutions, though some of the solutions may not be real and hence must be discarded. Perhaps unsurprisingly, one additional, i.e., seventh, interagent distance measurement is usually enough to disambiguate the multiple solutions, at least generically, so that a unique solution results.

A related result along these lines can be found in [15], expressed in terms of the merging of globally rigid frameworks. A globally rigid framework essentially is a framework of vertices and bars of known lengths joining the vertices, such that the specification of the lengths not only ensures the framework is rigid, but is in fact uniquely defined up to a congruence. Now, the coordinate-alignment problem of interest to us can generally be reformulated as one which 1) associates a 3-D (globally rigid) body with each agent (the body being a polyhedron with vertices defined by the agent positions at each time an inter-agent distance is measured), 2) associates a bar (or link between two bodies) with each interagent distance measurement, and 3) seeks conditions on the overall number of links and rules concerning the points on the bodies to which they can be incident to ensure that the entire framework is globally rigid, i.e., every joint or bar end in the framework has a computable location in a coordinate frame fixed to the framework, which cannot flex. The result states that if two globally rigid frameworks in an ambient 3-D space are joined by seven or more bars of fixed lengths (and the endpoints of the bars can only coincide to a limited extent, or are distinct), then the merged framework is again globally rigid. This means that if one of the frameworks before merging is localized, the whole framework after merging (and thus the second framework before merging) is localized.

The lower bound, seven, carries over to the localization problem. Furthermore, if all measurements were noiseless, there would be no reason to use more than seven measurements. Noise in practice is very likely, and so having more than seven measurements, in general, can be expected to reduce the localization error and, less obviously, to mitigate errors in timing synchronization. This is discussed much later, as also is the existence of exceptional trajectories (which can always be avoided) for which no solution can be found, irrespective of the number of measurements and solution methods.

III. SOLUTION TO THE TWO-AGENT PROBLEM

In this section, we focus on the two-agent case and, later note in a Remark, how the ideas can be applied to certain multiagent arrangements. We first present, in Section III-A, an SDP-based approach for solving a constrained least-squares problem, where no less than seven measurements are assumed available. This solution actually is a relaxed solution (in the sense of optimization theory) of the problem of interest, and so, a further step may have to be undertaken to deal with possible consequences of the relaxation. This further step may induce failure of the constraint that the matrix $R$ is orthogonal with determinant 1, and so, we show how to modify the outcome of the calculations of Section III-A to achieve an orthogonal matrix in Section III-B.

As we shall see later in this section, the SDP relaxation solves a least-squares error problem subject to a number of quadratic equality constraints, and therefore, any obtained optimal solution may not be optimal in a maximum likelihood sense. The reason is that the distance measurement errors are mixed and transformed nonlinearly during the process of deriving the least-squares error objective function. This holds true in the case where the measurement errors are zero-mean Gaussian. In the special case where the measurements are error-free, both optimization problems are equivalent in the sense that they produce the same optimal solutions. This point is brought into a sharper focus later in this paper when we introduce the index relevant to the MLE calculation. This index is evidently quite different from that of the SDP relaxation.

In Section III-C, we show how a maximum likelihood estimate of the matrix $R$ and vector $T$ defining the coordinate frame relation can be obtained by a gradient flow. The success of the gradient-based solution to converge to an MLE optimal solution, and not get trapped in a local minimum, requires that the initial search is started in the vicinity of an optimal solution. This is where the outcome of SDP relaxation analysis, developed in this section, becomes relevant. The gradient flow is nonstandard, in that the orthogonality constraint on the rotation matrix needs to be preserved by the flow. Section III-D reviews situations where agent trajectories are nongeneric, while Section III-E explains how to handle timing issues arising from lack of synchronization of clocks between the UAVs.

Consider having two agents, whose positions are $p_x = [x_1, x_2, x_3]^T$ and $p_y = [y_1, y_2, y_3]^T$ in their respective coordinate systems. Suppose further that agent $x$ is navigating in a global frame (e.g., has access to GPS coordinates). Agent $y$ may express its position in the global frame through the rigid transformation, $\tilde{p}_y = Rp_x + T$, where $R$ is a $3 \times 3$ rotation matrix and $T$ is a 3-vector. As before, $d$ denotes an interagent distance.
A. SDP Relaxation

Suppose $R = \{r_{ij}\}$, $T = [t_1, t_2, \ldots, t_N]^\top$. There holds for measurements at time $t_1, t_2, \ldots, t_N$ (with the time index suppressed in the following equations)

$$d^2 = \|\tilde{p}_y - p_x\|^2. \quad (2)$$

The equation for $d^2$ can be written as

$$d^2 = \|R \tilde{p}_y + T - p_x\|^2
= \|p_y\|^2 + \|T\|^2 + \|p_x\|^2 + 2T^\top R p_y - 2p_x^\top R p_y - 2p_x^\top T. \quad (3)$$

The fourth term involves quadratic expressions in the unknowns, i.e., the entries of $R$ and $T$. The fifth and sixth terms involve linear terms only. The end result is the following quadratic constraint on the entries of $R$ and $T$, presented to place all the unknowns on the right-hand side of the equation

$$d^2 - \|p_x\|^2 - \|p_y\|^2
= -2x_1y_1r_{11} - 2x_1y_2r_{12} - 2x_1y_3r_{13}
- 2x_2y_1r_{21} - 2x_2y_2r_{22} - 2x_2y_3r_{23}
- 2x_3y_1r_{31} - 2x_3y_2r_{32} - 2x_3y_3r_{33}
- 2x_1t_1 - 2x_2t_2 - 2x_3t_3
+ 2y_1 \sum r_{1i}t_i + 2y_2 \sum r_{2i}t_i + 2y_3 \sum r_{3i}t_i + \sum t_i^2. \quad (4)$$

If we regard each summand on the right-hand side of (4) as a product of known values $-2x_1y_1, -2x_1y_2, \ldots, 1$ and independent unknowns $r_{11}, r_{12}, \ldots, \sum t_i^2$, one can seek to solve the set of equations with a sufficient number of measurements. Generally, 16 measurements are sufficient because there are 16 linearly independent unknowns $r_{11}, r_{12}, \ldots, \sum t_i^2$ (at least if the associated coefficient matrix has full rank, which proves to be almost always the case). When the coefficient matrix is close to singular, more measurements are required. If there are more measurements than required, a least-squares solution is used.

In the above solution process, we (temporarily) regard each summand in (4) as a product of a known value and an independent unknown. This, therefore, leaves out the consideration of the nonlinear constraints on the unknowns (though recall that (4) only arises after use of a number of such constraints). These constraints, if accounted for, reduce the number of measurements required and improve the estimation accuracy in the presence of noise as we now argue in more detail.

Define $\Theta = [\theta_1, \theta_2, \ldots, \theta_{16}]^\top$ to be the 16-vector of unknowns $r_{11}, r_{12}, \ldots, \sum r_{3i}t_i, \sum t_i^2$ in (4) and $A[k] = [a_{k1}, a_{k2}, \ldots, a_{k16}]$ to be the row vector of known values $-2x_1y_1, -2x_1y_2, \ldots, 1$ in (4) at time $t = t_k$. Furthermore, assume for convenience that there are 16 measurements. Suppose $A = [A[1]^\top, A[2]^\top, \ldots, A[16]^\top]^\top$ and $b$ is a column vector, with the $k$th entry being $z[k]^2 - \|p_x[k]\|^2 - \|p_y[k]\|^2$, where $z[k]$ is the $k$th noisy measurement of $d$.

The optimization problem we are trying to solve is

$$\arg\min_{\Theta} \frac{1}{2} \|A\Theta - b\|^2$$
subject to $C(\Theta) = 0 \quad (5)$

where $C(\Theta)$ expresses all the constraints among the $\theta_i$. Note that there are altogether ten independent constraints expressed in $C(\Theta)$, as listed in Appendix I. Additional four constraints are listed, which are not independent of the ten just mentioned.

We now show how to solve (5), using SDP. Define

$$X = \begin{bmatrix} \Theta & \Theta \\ -1 & -1 \end{bmatrix}^\top.$$

Solving for $\Theta$ is equivalent to solving for $X$ with the constraints that

1) $X$ is positive semidefinite (denoted by $X \succeq 0$);
2) $\text{rank}(X) = 1$;
3) the bottom right corner element $X_{17,17} = 1$.

Suppose that $(u, v)$ denote the inner product of two matrices $u$ and $v$, i.e., $(u, v) = \text{trace}(u^\top v)$, and define $P = [A \ b]^\top [A \ b]$

The objective function $\frac{1}{2} \|A\Theta - b\|^2$ can be written as $\frac{1}{2} (P, X)$.

Similarly, the constraints in (19) can be written in terms of $X$ as

$$(Q_i, X) = q_i, \quad i = 1, \ldots, 10$$

where $Q_i$ are $17 \times 17$ symmetric matrices and $q_i$ are ten scalars. Now, the optimization problem becomes

$$\arg\min_{\Theta} \langle P, X \rangle$$
subject to $(Q_i, X) = q_i, \quad i = 1, \ldots, 10$

\[ X_{17,17} = 1 \]

\[ X \succeq 0 \]

\[ \text{rank}(X) = 1. \quad (6) \]

A naive SDP relaxation is given by dropping the rank constraint. After dropping the rank constraint, $C_i = 0, i = 1, \ldots, 10$, no longer imply $C_i = 0, i = 11, \ldots, 14$, so it is advisable to add these last four constraints as well. As observed in a number of simulations, the above additional constraints are helpful to obtain a low-rank solution for generic trajectories and, therefore, helpful to reduce the rounding error when we recover the rank-1 solution using singular value decomposition (SVD), as shown later in this paper. In addition, because the rotation matrix entries should satisfy $-1 \leq r_{ij} \leq 1$, one can also consider including reformulation-linearization-technique constraints [24] to further tighten the SDP relaxation.

After computing the SDP relaxation solution, we can find its best rank-1 approximation under matrix 2-norms by using SVD and setting all but the largest singular value to zero (see [25, Section 4.3]). In fact, it is generally a very accurate approximation because the solution of the relaxed problem is already close to rank 1. (In a number of numerical simulations, the largest singular value of the
solution turned out to be generally $10^2$–$10^5$ times larger than the second largest singular value.)

The solution of the above SDP can then be used as an initial condition to the gradient optimization presented below in Section III-C. One should note that the number of measurements used in the SDP method can increase arbitrarily and decrease to 7. Although the validity of the SDP approach does not straightforwardly imply that the minimum number of measurements is 7, we nevertheless assume that this constraint always holds; this will be discussed in more detail in the simulation section.

B. Obtaining a Rotation Matrix

In the noisy case, let $\tilde{T}$ and $\tilde{R}$ denote the estimated value of $T$ and $R$, respectively. Note that the imposition of the rank-1 constraint on the SDP matrix $X$ (from which entries of $\tilde{T}$ and $\tilde{R}$ can be determined) through approximation, as a final tidy-up step of the algorithm, may destroy the orthogonality of $\tilde{R}$, though up to that point, SDP guarantees orthogonality by virtue of the equality constraints. In this case, the obtained $\tilde{R}$ may not satisfy all the conditions to be a rotation matrix; therefore, one more step can be taken to find the rotation matrix $\bar{R}$ that has the closest Frobenius norm to $\tilde{R}$. Thus, one seeks

$$\bar{R} = \arg \min_{\Omega} \| \Omega - \tilde{R} \| \_F \quad \text{subject to} \quad \Omega \Omega^\top = I, \det \Omega = 1$$

where $\| \cdot \| \_F$ denotes the Frobenius norm.

This minimization problem is a special case of the orthogonal Procrustes problem [26, pp. 29–34]. To find this orthogonal matrix $\bar{R}$, the SVD is used

$$\bar{R} = U \Sigma V^\top.$$

Suppose $J$ is a diagonal matrix with the last entry on the diagonal being $-1$ and all other entries on the diagonal being 1. The solution of this constrained version of the orthogonal Procrustes problem is

$$\bar{R} = \begin{cases} 
UV^\top, & \text{if } \det(UV^\top) = 1 \\
UJ V^\top, & \text{if } \det(UV^\top) = -1.
\end{cases}$$

(9)

Note that $\det(\Sigma - I)$ can be used as an error measure.

C. Maximum Likelihood Estimation and Gradient Flow on a Manifold

In principle, in the presence of noise, the best choices for $R$ and $T$ are the maximum likelihood estimates of those quantities. As in many estimation problems, closed-form solution expressions rarely exist, and minimizing the performance index often leads to convergence to a local minimum. Our approach is to use the solution of the SDP relaxation to initialize a gradient descent algorithm. As examples later show, this proves highly effective.

The entries of $R$ and $T$ constitute the variables being estimated. A somewhat nonstandard descent algorithm (gradient descent on a manifold) is required to ensure that the orthogonality constraint is preserved.

Suppose the measurement distance, $z[k]$, is contaminated by a Gaussian noise with zero mean and standard deviation $\sigma$, i.e., the measurement sensor delivers $\tilde{z}[k] = z[k] + \xi, \xi \sim N(0, \sigma^2)$. We assume in this section that $p_x$ and $p_y$ can be obtained without noise, and that the measurement noise values at different times are independent.

Now, we obtain $\tilde{z}[k] - \| Rp_x[k] + T - p_x[k] \| \sim N(0, \sigma^2)$. The likelihood function is

$$\mathcal{L}(p_x, p_y, z| R, T) = \frac{1}{\sigma \sqrt{2\pi}} \prod_{k=1}^{N} \exp \left[-\frac{(\tilde{z}[k] - \| Rp_x[k] + T - p_x[k] \|^2)}{2\sigma^2} \right]$$

or equivalently

$$\log \mathcal{L}(p_x, p_y, z| R, T) = \sum_{k=1}^{N} \left[-\frac{(\tilde{z}[k] - \| Rp_x[k] + T - p_x[k] \|^2)}{2\sigma^2} \right] - \log(\sigma \sqrt{2\pi}).$$

Therefore, the maximum likelihood estimate is given by solving the optimization problem

$$\bar{R}, \bar{T} = \arg \min_{R,T} \| R \tilde{p}_x[k] + T - p_x[k] \| \sum_{k=1}^{N} \left[-\frac{(\tilde{z}[k] - \| Rp_x[k] + T - p_x[k] \|^2)}{2\sigma^2} \right]$$

subject to $RR^\top = I$ \det(R) = 1.$$

As in many MLE estimation problems, the index is not convex, and minimization is not necessarily straightforward. A particular problem with using gradient descent on any nonconvex function is to find an initialization within the capture region of the global minimum, and this is where the calculations of the previous subsections become relevant, if not critical. We use the result from linear processing (with the number of measurements being 16) or preferably the SDP approach of the previous subsections (with the number of measurements being greater than or equal to 7 and the solution adjusted if necessary via the Procrustes algorithm) to initialize a gradient descent algorithm aimed at finding the minimum.$^2$

It is useful for this purpose to know how to compute the gradient of a function of a special orthogonal matrix on the manifold of special orthogonal matrices. Consider $f: SO_3 \rightarrow \mathbb{R}$, mapping special orthogonal matrices to the reals. Suppose we want to compute the gradient, reflecting the orthogonal property.

The general idea (technically a consequence of the fact that $SO_3$ is a Riemannian manifold and so inherits a metric

---

$^2$In recently published work [27] studying autonomous underwater vehicle localization using distance-only measurements to a non-GPS-equipped vehicle, minimization of the same MLE index is tackled using a form of approximation for the index, with the application of a parallel projection algorithm. Proper comparison with the methods of this paper cannot be made, due to the limited details in this paper. How issues of initialization can be effectively tackled is also not clear.
from the Euclidean space in which it is embedded [28]) is: first, we consider a point \( R \in \mathbb{R}^{3 \times 3} \) on the \( SO_3 \) manifold and compute the gradient \( \frac{df}{\theta R} \) in the standard way; then, we project it onto the tangent space of \( SO_3 \) at the point \( R \). For this purpose, we need to have the tangent space and normal space of \( SO_3 \) at some point \( R \), and also the projection of a vector to the tangent space.

**Lemma 1:** i) The tangent space of \( SO_3 \) at \( R \in SO_3 \) is the set of \( \mathcal{P} \) such that [28]

\[
\mathcal{P}^T R + R^T \mathcal{P} = 0
\]

or equivalently

\[
T_{SO_3(R)} = \{ RQ, Q + Q^T = 0 \}
\]

and the normal space is

\[
N_{SO_3(R)} = \{ RS, S - S^T = 0 \}
\]

where \( T_{SO_3(R)} \) and \( N_{SO_3(R)} \) denote the tangent and normal spaces at \( R \), respectively.

ii) Furthermore, suppose at a point \( R \in SO_3 \), the gradient of \( f \) in \( \mathbb{R}^{3 \times 3} \) is

\[
\frac{\partial f}{\partial R} = M.
\]

Then, the projection of \( M \) on the tangent space \( T_{SO_3(R)} \) is given by

\[
M_T = \frac{1}{2} M - \frac{1}{2} RM^T R.
\]

**Proof:** The proof is given in Appendix II.

Now, suppose \( f = \sum_{i=1}^{N} \|z[k_i] - ||R_{p_i}[k_i] + T - p_i, \| k_i \|^2 \). It is straightforward to obtain the gradient \( M_T \) and \( \frac{\partial f}{\partial T} \), and so, we have the gradient flow

\[
\dot{R} = -M_T
\]

\[
\dot{T} = -\frac{\partial f}{\partial T}.
\]

(13)

In summary, the SDP relaxation of Section III is likely to give a good initial condition, and then, a gradient method using a discretization of (13) is sufficient for solving this optimization problem. The Procrustes problem algorithm can be applied in each step to correct departure from orthogonality due to round-off and discretization error.

**Remark 1:** In the case of multiagent localization, the same algorithm can be adopted for those networks where each non-GPS equipped agent has at least seven measurements to a GPS-equipped agent. Thus, if there is only one GPS-equipped agent, the graph will be a star form; each of the GPS-denied agents (with INS) must have seven or more distance measurements between it and the GPS-equipped agent. This allows a collection of two-agent problems to be solved, possibly in parallel. Note that this topology may well correspond to the physical situation of one agent remaining high above a building canyon, while other agents navigate within the building canyon.

**Remark 2:** The objective function of the constrained linear least-squares problem and the objective function of the gradient descent method are different. The first one is designed to exploit as far as possible the linear occurrence of unknowns in the \( z_i \) expressions, while the second one is derived from a maximum likelihood estimator. As noted, the constrained least-squares problem should be used to find an appropriate starting point for the gradient descent method, which then finds the maximum likelihood estimate.

**D. Observability Issues and Nongeneric Trajectories**

With certain nongeneric trajectories, localization is not achievable by taking only distance measurements. To understand this phenomenon, we first study the linear estimator, as follows. In (5), if we have more than 16 distance measurements, we can disregard the constraints and obtain an estimate by taking the least-squares solution \( \theta = (A^T A)^{-1} A^T b \). However, if the pair of agents fly along a trajectory such that \( A \) is singular, a solution cannot be found. Nevertheless, these problems can be avoided if the agents’ trajectories contain small deviations, either intended (e.g., part of the path plan or maneuvers) or unintended (e.g., wind gusts). As such, it is safe to assume that a generic trajectory would not be problematic.

Similarly, when applying the SDP or MLE approach, one could face a similar singularity problem. The cases of singular trajectories include, but are not restricted to the following: 1) each agent flies in an affine subspace of \( \mathbb{R}^3 \) (i.e., their trajectories are individually coplanar); or 2) the distance between agents remains constant at all time. Again, these problems can be avoided if agents’ trajectories have some oscillations.

**E. Unsynchronized Clocks and Distance Measurements Bias**

The most commonly used distance measurement method for UAVs is time of arrival (ToA). The ToA ranging technique usually requires that the transmitter and receiver have synchronized clocks. With synchronized clocks, the receiver can compute the signal transmission time and obtain a distance measurement. In the case of unsynchronized clocks, constant measurement bias may be added to the computed distances and degrade the performance of a localization system.

To tackle the problem with biased distance measurements, we consider an initially unknown constant \( s \) being added to our measurements,\(^3\) i.e., the obtained distance measurement is \( d_o \), but the actual distance becomes \( d_o + s \). We seek to estimate the unknown \( s \) as well as the transformation between the two coordinate frames. Making an

---

\(^3\) The assumption of a constant clock bias is essential to our solution approach, but may not apply in general. Small UAVs with cheap oscillators may have a drifting bias during the measurement interval that causes a problem with localization.
obvious change to (4), we obtain

\[ d_o^2 - \|p_o\|^2 - \|p_s\|^2 = -2x_1y_1r_{11} - 2x_1y_2r_{12} - 2x_1y_3r_{13} - 2x_2y_1r_{21} - 2x_2y_2r_{22} - 2x_2y_3r_{23} - 2x_3y_1r_{31} - 2x_3y_2r_{32} - 2x_3y_3r_{33} - 2x_1t_1 - 2x_2t_2 - 2x_3t_3 + 2y_1 \sum r_{11}t_1 + 2y_2 \sum r_{22}t_2 + 2y_3 \sum r_{33}t_3 + \sum t_i^2 + s^2 + 2d_0s. \] (14)

Now, the vector of unknowns \( \Theta = [\theta_1, \theta_2, \ldots, \theta_{17}]^T \) is a 17-vector of unknowns \( r_{11}, r_{12}, \ldots, \sum r_{33}t_3, \sum t_i^2 + s^2 \), and \( s \) in (14). Following a very similar solution procedure to that of Section III-A (either linear processing or SDP, with the minimum number of measurements being increased by 1), one can obtain estimates of \( R, T, \) and \( s \). If SDP is the intended solution method, constraints \( C_9 \) and \( C_{10} \) in Appendix I should also change to

\[ C_9' = \theta_1^2 + \theta_2^2 + \theta_3^2 + \theta_4^2 - \theta_6 = 0 \]
\[ C_{10}' = \theta_1^2 + \theta_4^2 + \theta_5^2 + \theta_7^2 - \theta_6 = 0. \] (15)

IV. SIMULATIONS, PERFORMANCE EVALUATION, AND FLIGHT DATA EXAMPLE

A. Performance Metric

In order to evaluate the performance of our localization algorithm, a performance metric is selected in this subsection. As we have already defined in the beginning of Section III, \( p_x \) and \( p_y \) denote the coordinates of two agents \( x \) and \( y \) in their own local coordinate systems. Measurements are taken at time instances \( \tau = \tau_k, k = 1, \ldots, N \). Our objective is to estimate \( R_{p_x}(\tau) + T - p_s(\tau) \), the relative position of agent \( y \) in agent \( x \)'s coordinate system, when the last distance measurement \( d(\tau_N) \) is taken.

Without loss of generality, we may assume that each agent is located at the origin of its respective local coordinate system when the last measurement is taken, i.e., \( p_x(\tau_N) = 0 \). This can be done without loss of generality because the origin of a local coordinate system can be selected at our convenience. Now, our localization objective, the relative position of agent \( y \) in agent \( x \)'s coordinate system, is given by

\[ R_{p_s}(\tau) + T - p_s(\tau_N) = T. \] (16)

That is to say, our objective is to obtain accurate estimates of \( T \). For this purpose, the error in the estimate of \( R \) is irrelevant.

Also note that

\[ d(\tau_N) = \|R_{p_s}(\tau_N) + T - p_s(\tau_N)\| = \|T\| \] (17)
i.e., the norm of \( T \) is directly measured by \( d(\tau_N) \). Hence, our performance metric should be a metric of directional error of \( T \). A metric of directional error is defined as

\[ E_{\text{direction}} = \arccos \left( \frac{\tilde{T}^\top T}{\|\tilde{T}\|\|T\|} \right) \] (18)

where \( T \) is the true translation vector and \( \tilde{T} \) is its estimated counterpart.

Other metrics could be considered. However, the primary goal of this paper is the localization of a peer agent. For measuring the error, evaluation of \( E_{\text{direction}} \) is enough. In addition to localization, if aligning coordinate systems is also a requirement (and this is frequently so), then we have to look at the full attitude error (in terms of \( R \)). Due to space limitation, the analytic evaluations of the full attitude error are not provided, but simulation data of the total angular error are presented below. Generally speaking, less error in \( T \) means less error in \( R \), as shown in a number of simulations.

B. Simulation and Comparison Between SDP and Gradient Descent

In the simulations below, white Gaussian noise is added to interagent distance measurements, and there are seven or more such measurements. Furthermore, the SNR denotes the signal-to-noise ratio in dB, defined as \( \text{SNR}_{\text{dB}} = 20 \log_{10} \left( \frac{A_{\text{signal}}}{A_{\text{noise}}} \right) \), where \( A_{\text{signal}} \) is the average distance between agents and \( A_{\text{noise}} \) is the root-mean-square amplitude of the noise added to the distance measurements.

1) In each figure, we show the directional error of \( T \) defined in (18) in radians versus number of distance measurements.
2) Fig. 2 shows a typical example as used in the simulations illustrating the actual and estimated trajectory of the non-GPS equipped vehicle.
3) Fig. 3 depicts the noiseless case, i.e., \( \text{SNR} \to \infty \). Similarly, Figs. 4–6 depict the cases where \( \text{SNR} = 30, \text{SNR} = 20, \) and \( \text{SNR} = 10 \), respectively. It is well known that, in practice, the distance accuracy (1-sigma distance error) is inversely proportional to signal bandwidth as well as the square root of the
For passive RF detection problems, the SNR in free space is inversely proportional to distance squared. At 1 km, the SNR is usually high, say 20 dB or more. On the other hand, the INS that provides local coordinates is assumed to be driftless over the time duration of \( n \) measurements. This assumption is reasonable because at least for the UAVs we used in field testing, the INS achieves an accuracy of 0.2% of the distance traveled. Obviously, if the INS does incur a significant drift, the performance of our estimation method will degrade. Clearly, this peer localization method is suitable for vehicles employing higher grade inertial navigation systems. Such vehicles include high-value platforms, long-range missiles, etc. It is important to note that the drift is only required to be insignificant within a short period of time.

4) Gradient descent refinement can provide an improved result in comparison to SDP relaxation. This is partly because gradient descent refinement uses a better estimator, and partly because of the relaxation error in SDP. However, it is also notable that SDP relaxation is crucial in providing the initial condition to avoid convergence to a local minimum.

5) As noted in the introductory section, the paper [13] studies a similar problem to ours, with application in robotics systems. Despite the different application scenarios, we implemented the algorithm of [13] in the context of our application and made a direct comparison of the directional error of \( T \) to benchmark the performance of our algorithm.

6) From [12], one notes that with six distance measurements, the 3-D geolocation problem has 40 solutions, where some of the solutions may not be real and must be discarded. The paper [15] further shows that in the absence of noise, one additional measurement can almost always disambiguate the solutions and with
seven distance-only measurements, a unique solution can be found. Although the use of SDP followed by the gradient-based MLE algorithm as proposed in this paper apparently does not require a minimum number of measurements, an implicit assumption is that the number of measurements is indeed greater than or equal to 7.\(^4\) On the other hand, the algorithm in [13] requires at least ten measurements, and even with ten measurements, the algorithm sometimes fails to converge. In our implementation, we find that the algorithm in [13] can reliably converge with 12 or more measurements. Hence, we record the error data of that algorithm starting from 12 measurements.

Figs. 3–6 show the change of direction error in degrees as the number of measurements increases. In these figures, the blue curve shows the result of the algebraic method in [13], the red curve shows the result of SDP, and the yellow curve shows the result of SDP with gradient descent refinement. With each number of measurements and each level of SNR, each point plotted in the above figures is the average of 200 simulations with random vehicle trajectories (a pure random walk, whose derivative is Gaussian distributed with zero mean and a 1-m standard deviation in all directions), and measurements are taken every step of the random walk. In each simulation, we resample both the random walk and measurement noise. We have left \( R \) and \( T \) unaltered because the results are not sensitive to the choice of \( R \) and \( T \). We choose \( T = [0.4 \; 0.8 \; 1]^T \) and

\[
R = \begin{bmatrix}
-0.22 & -0.35 & -0.90 \\
0.90 & -0.42 & 0.06 \\
-0.37 & -0.83 & 0.41
\end{bmatrix}.
\]

\( R \) is generated using Euler angles (in radians) [1 2 3].

Evidently, our estimator is not a linear estimator; this means that (as is common for nonlinear estimators) it is not necessarily unbiased, and in fact, we have found that it is biased. Therefore, as the number of useful measurements increases, the direction error does not converge to zero in Figs. 4–6. The high-level conclusion is that it is possible to handle a large number of measurements, and there is graceful degradation of estimation quality with increasing noise, with no apparent collapse at a certain point.

Compared to the method provided in [13], our approach in general achieves better accuracy, especially under excessive noise. The SDP processing by itself is better in all cases with noise, and the performance of SDP with gradient refinement is superior in all cases. Most importantly, our method is far more robust in three aspects. First, when the noise is large, the method proposed in [13] often diverges (the highest percentage of diverging cases reported is 65%, corresponding to odometer error \( \sigma_o = 0.1 \) m). Our method, on the other hand, despite exhibiting gradually degrading accuracy, never has this divergence issue. Furthermore, the proposed method in [13] only works with a minimum of ten distance measurements. In the case of seven, eight, or nine measurements, our method will still work, but the one in [13] will not. Finally, our proposed SDP+MLE framework can handle system bias with only minor modifications. For example, Section III-E proposes minor algorithm modifications when there is measurement bias due to unsynchronized clocks.

Both the quaternion method in [13] and our method relies on gradient methods (e.g., interior-point method) to find solutions. Note the SDP relaxation is a way to convert nonconvex optimization problems into convex optimization problems, and the solution method is the same with any standard convex optimization problems. As a result, the computational complexity is not drastically different. We do not have an opportunity to test the method in the onboard computational devices on a UAV, but the SDP+MLE method finishes within 1 s on a laptop. In general, our method may use more computational time than [13], but this should not be perceived as a major limitation.

### C. Trial on Real Data

In order to validate the key assumptions in our model (including measurement noise level, odometer accuracy, measurement frequency, etc.), we applied our SDP+MLE method on real data. In particular, real flight data provided by the Australian Defence Science and Technology Group is used in this section to evaluate the performance of the proposed method in practice. The data consists of the true positions of UAV1 in global coordinates, the positions of UAV2 in its local INS coordinates, and the distance measurements between the pair of agents. The relevant numbers are summarized in Table I.

<table>
<thead>
<tr>
<th>Time(s)</th>
<th>UAV1 global Coordinates</th>
<th>UAV2 Local Coordinates</th>
<th>Distances (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>x (m)</td>
<td>y (m)</td>
<td>z (m)</td>
</tr>
<tr>
<td>3.2</td>
<td>349.1</td>
<td>-924.1</td>
<td>374.4</td>
</tr>
<tr>
<td>12.9</td>
<td>578.2</td>
<td>-945.0</td>
<td>371.9</td>
</tr>
<tr>
<td>22.7</td>
<td>781.0</td>
<td>-970.3</td>
<td>372.3</td>
</tr>
<tr>
<td>32.5</td>
<td>936.8</td>
<td>-715.4</td>
<td>373.7</td>
</tr>
<tr>
<td>42.3</td>
<td>1007.9</td>
<td>-522.7</td>
<td>373.3</td>
</tr>
<tr>
<td>52.1</td>
<td>992.0</td>
<td>-329.4</td>
<td>373.3</td>
</tr>
<tr>
<td>61.8</td>
<td>889.8</td>
<td>-191.0</td>
<td>373.2</td>
</tr>
<tr>
<td>71.6</td>
<td>660.1</td>
<td>38.6</td>
<td>372.9</td>
</tr>
<tr>
<td>80.5</td>
<td>431.4</td>
<td>36.6</td>
<td>373.1</td>
</tr>
<tr>
<td>91.1</td>
<td>189.7</td>
<td>-94.9</td>
<td>373.3</td>
</tr>
<tr>
<td>100.9</td>
<td>33.9</td>
<td>262.2</td>
<td>373.6</td>
</tr>
</tbody>
</table>

In the case of less than or equal to five measurements, the gradient of the objective function of the MLE optimization is always zero on a manifold, and the final result is randomly located on the manifold. In addition, in the case of six measurements, there is no way to deal with the issue of disambiguating local optima.

\( ^4 \)In the case of less than or equal to five measurements, the gradient of the objective function of the MLE optimization is always zero on a manifold, and the final result is randomly located on the manifold. In addition, in the case of six measurements, there is no way to deal with the issue of disambiguating local optima.
Fig. 7. 3-D plot of the UAV positions in global coordinate systems.

5 min without an air data unit). This all means that, in our field test data, the global coordinate data of UAV 1 can be regarded as error-free, and the local coordinate data of UAV 2 are obtained from an IMU with an air data unit (Spatial Dual from Advanced Navigation). Furthermore, the error for distance measurements is about 15 m, while UAVs are on average 1500 m apart from each other, i.e., the SNR is roughly 40 dB.

Here, distance measurements are postprocessed distance measurements created by adding noise to the distance obtained from GPS. In this computation with real data, we take UAV2’s perspective, i.e., it obtains UAV1’s global coordinates at several time instants through the processing of distance measurements via ToA of those signals (in this experiment, both UAVs’ clocks are accurately synchronized), in conjunction with own IMU acceleration and angle rate data. Feeding those data as inputs, UAV2 computes its own positions in the global coordinates at those time instants using SDP followed by gradient MLE. In Fig. 7, circles denote the true positions of UAV 1 in global coordinates and triangles denote the computed positions of UAV 2 in the global coordinates. For comparison purposes, the true trajectory of UAV2 in the global coordinates is also directly measured and is displayed with a solid line. If we compute the error metric defined in (18), we find that the direction error of our estimation is 0.03 rad. This result is consistent with the simulation result with 11 measurements and an SNR of 30 dB.

In Table II, we see that the localization algorithm achieves better accuracy in the north and east directions but poorer accuracy in height. This can be explained by the observation that the trajectories of the UAVs are very close to coplanar, and that the distance measurements shed very little information on the height dimension. Indeed, if we change the height recorded by a UAV’s inertial sensor by a few meters, a “mirror solution” will be obtained. In fact, if the flying trajectories are exactly planar, then as remarked earlier, it is not possible to recover bearing information from distance-only measurements due to singularity issues. However, even with this close-to-singular setting, our method successfully recovered location information with reasonable accuracy.

The data provided include 11 distance measurements. Although the result shown in Fig. 7 uses all 11 measurements, we only need seven measurements to recover the positions of UAV2 in the global coordinates. Further calculations show that results obtained with fewer distance measurements have progressively decreasing accuracy as expected. Actually, this field test perhaps delivers good results because the two UAVs are close to each other, viz., the two UAVs are 1 km apart on average and move an average of 0.2 km between taking distance measurements. Our method performs well in this setting with the above two numbers on the same scale. In real applications, however, if the two UAVs are very far apart from each other, the accuracy of localization degrades. In order to improve localization, the “agent geometry and mobility” might have to be improved during distance measurement collection. This may mean that UAVs have to travel much faster and occasionally execute direction changes.

V. CONCLUSION

In this paper, we first proposed a novel semidefinite optimization approach for solving the problem of 3-D mobile localization of a GPS-denied agent using distance-only measurements. After that, a maximum likelihood estimator is used in a further approach to enhance the accuracy of localization, with simulations using real field test data.

Future work includes introducing a systematic treatment of the multiagent case, possibly including batch approaches, using a similar procedure to that of this paper. We are also involved in a separate study of localization of GPS-denied agents using bearing-only (azimuth and elevation) measurements. Furthermore, finding optimal trajectory for UAV-to-UAV distance-based positioning is also a future research direction as UAV trajectories have a significant effect on the accuracy of such position methods.
Because our estimator is not linear, bias will be present in our final estimates. Another direction for future research is to provide a bias correction procedure for both the pair-agent and multiagent cases.

APPENDIX I

The ten constraints are listed as follows (the first six coming from orthogonality alone). Let \( C_i(\theta) \) be the \( i \)th constraint; we have

\[
\begin{align*}
C_1 &= \theta_1^2 + \theta_2^2 + \theta_3^2 - 1 = 0 \\
C_2 &= \theta_2^2 + \theta_3^2 + \theta_4^2 - 1 = 0 \\
C_3 &= \theta_3^2 + \theta_4^2 + \theta_5^2 - 1 = 0 \\
C_4 &= \theta_4^2 + \theta_5^2 + \theta_6^2 - 1 = 0 \\
C_5 &= \theta_5^2 + \theta_6^2 + \theta_7^2 - 1 = 0 \\
C_6 &= \theta_6^2 + \theta_7^2 + \theta_8^2 - 1 = 0 \\
C_7 &= \theta_7^2 + \theta_8^2 + \theta_9^2 - 1 = 0 \\
C_8 &= \theta_8^2 + \theta_9^2 + \theta_{10}^2 - 1 = 0 \\
C_9 &= \theta_9^2 + \theta_{10}^2 + \theta_{11}^2 - 1 = 0 \\
C_{10} &= \theta_{10}^2 + \theta_{11}^2 + \theta_{12}^2 - 1 = 0.
\end{align*}
\]

(19)

Note there are ten independent equality constraints and 16 independent variables, so the problem has six degrees of freedom. That is consistent with the fact that each of the rotation matrix \( R \) and the translation matrix \( T \) has three degrees of freedom. One should also note that the equation set used to express those constraints is not unique. In fact, there are four other constraints being dropped here, which can be derived from \( C_i = 0, \ i = 1, \ldots, 10 \). The additional constraints, of which the first three come from orthogonality of \( R \), are

\[
\begin{align*}
C_{11} &= \theta_1^2 + \theta_2^2 + \theta_3^2 - 1 = 0 \\
C_{12} &= \theta_1^2 + \theta_2^2 + \theta_3^2 - 1 = 0 \\
C_{13} &= \theta_1^2 + \theta_2^2 + \theta_3^2 - 1 = 0 \\
C_{14} &= \theta_1^2 + \theta_2^2 + \theta_3^2 - 1 = 0.
\end{align*}
\]

During the variable lifting process required in the SDP relaxation, these four constraints can no longer be inferred from the ten constraints above and have to be accounted for in the relaxation problem.

APPENDIX II

This appendix presents the proof of Lemma 1.

**Proof:** The first part of the proof regarding the tangent and normal spaces of \( SO_3 \) (i) is given in [28], and we simply give an outline here. Any infinitesimal perturbation \( \delta R \) of an orthogonal matrix \( R \), which preserves orthogonality, i.e., \( R^\top R = I \), necessarily satisfies \( R^\top \delta R + (\delta R)^\top R = 0 \). This identifies the set \( P \), and the tangent space description is immediate. Furthermore, if \( N \) is a matrix in the normal space, there must hold by definition

\[
\text{trace}(N^\top P) = 0 \ \forall P \in T_{SO_3}(R).
\]

One can readily check that any matrix of the set \( N_{SO_3}(R) \) satisfies this requirement, and with slightly more work, that this set exhausts all such matrices.

We now prove that (ii) the projection of \( M \) on \( T_{SO_3}(R) \) is given by \( M_T = \frac{1}{2} M - \frac{1}{2} R M^\top R \).

First, let \( M_N = \frac{1}{2} M - \frac{1}{2} R M^\top R \). Observe that

\[
R^\top M_N = \frac{1}{2} R^\top M + \frac{1}{2} M^\top R
\]

is symmetric; therefore, \( M_N \in N_{SO_3}(R) \) is normal to the tangent space \( T_{SO_3}(R) \) at \( R \). Furthermore, we have

\[
R^\top M_T = \frac{1}{2} R^\top M - \frac{1}{2} M^\top R
\]

is skew symmetric; thus, \( M_T \in T_{SO_3}(R) \). Furthermore, because \( M_T + M_N = M \), \( M_T \) is the projection of \( M \) on \( T_{SO_3}(R) \).

\[\square\]

**REFERENCES**


Wireless sensor network localization techniques


A multi-agent approach to power system restoration


Simultaneous velocity and position estimation via distance-only measurements with application to multi-agent system control


Reaching a consensus in a dynamically changing environment: A graphical approach


Localization in three-dimensional wireless sensor networks

In The Art of Wireless Sensor Networks, H. M. Ammari, Ed.


Formation control using range-only measurements


On Fisher information reduction for range-only localization with imperfect detection


[8] L. Taff

Target localization from bearings-only observations


Bearing-only measurement self-localization, velocity consensus and formation control


Cooperative localisation of a GPS-denied UAV in 3-dimensional space using direction of arrival measurements


Translational velocity consensus using distance-only measurements

Bomin Jiang received the B.E. degree (First Class Hons.) in mechatronic systems engineering from The Australian National University, Canberra ACT, Australia, in 2014. He is working toward the Ph.D. degree with the Institute for Data, Systems, and Society, Massachusetts Institute of Technology, Cambridge, MA, USA.

In August 2014, he was awarded with a University Medal from the Australian National University upon completing his B.E. degree. In the Australian National University, the University Medals are not awarded in every discipline every year; they are reserved only for students who complete Honours Degrees with outstanding achievements in both coursework and research. In September 2017, he was awarded with the MIT Presidential Fellowship. His current research interests include networked dynamical systems, game theory, and econometrics.
Brian D. O. Anderson (M’66–SM’74–F’75–LF’07) was born in Sydney, NSW, Australia. He received the B.Sc. degree in mathematics and the B.E. degree in electrical engineering from the University of Sydney, Sydney, in 1962 and 1964, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1966.

He is an Emeritus Professor with the Australian National University, Canberra, ACT, Australia (having retired as a Distinguished Professor in 2016), a Distinguished Professor with Hangzhou Dianzi University, Hangzhou, China, and a Distinguished Researcher with National ICT Australia, Canberra. His current research interests include distributed control, social networks, and econometric modeling.

Dr. Anderson was the recipient of the IEEE Control Systems Award of 1997, the 2001 IEEE James H Mulligan, Jr., Education Medal, and the Bode Prize of the IEEE Control System Society in 1992, as well as several IEEE and other best paper prizes. He is a Fellow of the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, and the Royal Society, and a foreign member of the U.S. National Academy of Engineering. He holds honorary doctorates from a number of universities, including Université Catholique de Louvain, Belgium, and ETH, Zürich. He is a past president of the International Federation of Automatic Control and the Australian Academy of Science.

Hatem Hmam received the Ph.D. degree in electrical engineering from the University of Cincinnati, Cincinnati, OH, USA, in 1992.

In Late 1996, he joined the Defence Science and Technology Group, Edinburgh, SA, Australia, where he is currently a Senior Research Scientist with the Cyber and Electronic Warfare Division. His current research interests include signal processing, alternative positioning and navigation, and optimal sensor placement.