SIAM J. Control Optim.
Vol. 53, No. 6, pp. 3534–3561

SENSOR NETWORK LOCALIZATION ON THE GROUP OF THREE-DIMENSIONAL DISPLACEMENTS*

J. R. PETERS†, D. BORRA‡, B. E. PADEN†, and F. BULLO†

Abstract. We consider the problem of estimating relative configurations of nodes in a sensor network based on noisy measurements. By exploiting the cyclic constraints induced by the sensing topology to the network, we derive a constrained optimization on $SE(3)^n$. For the case of a network with a single cyclic constraint, we present a closed-form solution. We show that, in certain cases, namely restriction to pure rotation and pure translation, this solution is independent of the particular representation of the constraint function and is the unique, constrained minimizer of an appropriate cost. For sensing topologies represented by a general, weakly connected digraph, we present a solution method which is based on the limits of the solution curves of a continuous-time ordinary differential equation. We show that solutions obtained by our method satisfy all semicycle (generalized cycle) constraints induced by the sensing topology of the network. Further, we show through numerical simulation that for "Gaussian-like" noise models with small variance, our solution achieves noise reduction comparable to that of the least squares estimator for an analogous linear problem.

Key words. sensor networks, localization, nonlinear optimization, Euclidean group, Lie groups, displacements, configurations, gradient flows, manifolds

AMS subject classifications. 22E15, 34A34, 37C10, 37N35, 70B10, 70B15, 70Q05, 70G65

DOI. 10.1137/140957743

1. Introduction.

1.1. Problem description. Sensor localization is an important component in many network applications, as it provides the basis for more complex operations and computation [39]. Problems that require localization include the hand-eye coordination problem in robotics [15], the structure from motion problem in computer vision [22], and the camera calibration problem in visual sensing [18]. Accurate localization is crucial, as errors in configuration measurements often get amplified during other calculations. For example, in camera networks, small errors in calibration get amplified when imaging objects that are far from the camera's physical location. In many applications, it is only possible to obtain noisy, relative configuration measurements. Therefore, it is necessary to develop a methodology for estimating true relative measurements in order to guarantee a global configuration estimate that is both optimal in some sense and consistent with constraints induced by the network topology.

We consider a three-dimensional sensor network whose sensing capabilities can be represented by a directed graph. We assume that each sensor obtains noisy, relative configuration measurements of its neighboring sensors, and thus we associate to each edge in the sensing graph a displacement matrix, i.e., an element of $SE(3)$. An example of this type of sensor network and its associated sensing graph is shown in Figure 1. The composition of the true displacement matrices associated to edges

*Received by the editors February 19, 2014; accepted for publication (in revised form) September 14, 2015; published electronically December 8, 2015. This material is based upon work supported in part by NSF Award 1035917 and in part by ARO Award W911NF-11-1-0092.
http://www.siam.org/journals/sicon/53-6/95774.html

†Department of Mechanical Engineering and the Center for Control, Dynamical Systems and Computation, University of California at Santa Barbara, Santa Barbara, CA 93106 (jrpeters@engineering.ucsb.edu, paden@engineering.ucsb.edu, bullo@engineering.ucsb.edu).
‡Dipartimento di Scienze Matematiche, Politecnico di Torino, 10129 Torino, Italy (domenica.borra@polito.it).

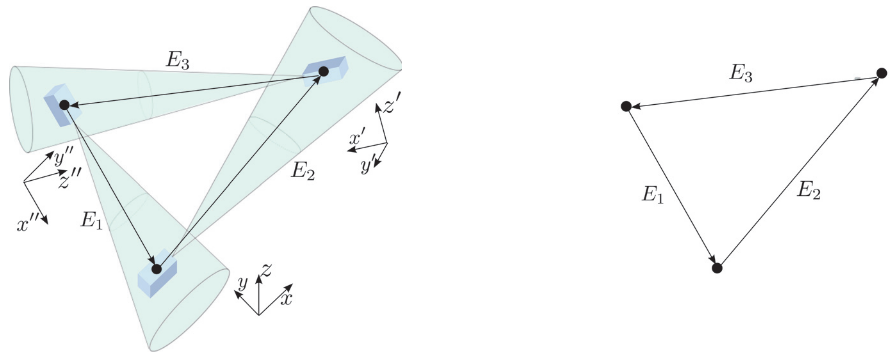

FIG. 1. *An example of a sensor network and its associated sensing graph.*

that form a semicycle (generalized cycle) must be equal to the identity, and thus for each semicycle we have an induced constraint. We seek to find relative displacements that minimize the least squares distance from the noisy measurements and satisfy all induced semicycle constraints.

1.2. Related work. Of relevance to this work is research concerning network localization in sensor networks, orientation and pose estimation in robotic sensor networks, properties of matrix Lie groups, namely $SO(3)$ and $SE(3)$, estimation and optimization on manifolds, and properties of cycle spaces in graphs.

There has been significant research on estimating node positions based on noisy angle of arrival, time of arrival, received signal strength, and distance-based measurements [28, 31]. The relative position localization problem is studied in [16, 32, 29], and the effect of network properties on the estimation error with extensive analytic results is investigated in [8, 9]. Some works, such as [11, 14, 28, 7], seek to find global position information through the use of anchor nodes with known locations. Other works, such as [40, 29], explore anchor-free methods for position localization. There has also been research which seeks to find global or relative configurations of sensor nodes by assigning to each an element of $SO(3)$ or $SE(3)$. In [36, 25], the authors use an unconstrained nonlinear optimization to derive a solution. In [8, 34], the authors study the problem of pose estimation from relative measurements in a distributed robotic sensor network. In [33] the authors consider the problem of orientation localization on $SO(2)$ under cyclic constraints induced by the network. We note that our work is largely an extension of the work in [33] to $SE(3)$.

There has been extensive research concerning matrix Lie groups and their properties, namely those properties relevant to estimation. A study of metrics on $SO(3)$ is presented in [26]. A survey of the state-of-the-art techniques used for rotation averaging is provided in [24]. Metrics and geodesic curves on $SE(3)$ are presented in [41]. Additional properties of $SE(3)$ are presented in the context of rigid body control in [13]. With regard to estimation on Riemannian manifolds, [19] studies the problem of geodesic regression and unconstrained least squares. In [2, 4], the authors formulate a nonlinear optimization and use a generalized line search technique in order to find iterates which converge to a local minimum. In [3], the authors provide some tools for constructing projection-like retractions on general manifolds.

There have also been authors who have studied the properties of cycle spaces on graphs. In [27, 37], the authors provide an extensive survey of the properties of cycle bases, including applications and algorithms for finding cycle bases. In [21, 6], the authors explore the problem of finding cycle bases with minimum total weight.

The problem considered in the present work is most closely related to those considered in [38, 35]. In [38], the authors seek to estimate absolute camera poses in $SE(3)$ based on noisy relative measurements. The authors construct a metric on the product space $SO(3) \times \mathbb{R}^3$ and present a gradient descent algorithm, which they show to converge on connected, symmetric sensing graphs. In [35], the authors present a discrete-time, rotation averaging based algorithm for estimating relative poses in the context of multiview image registration. The authors construct metrics based on rotational error and the relative translational location of a point in an absolute reference frame, explore optimality properties for a single cycle graph (focusing on rotation), and prove convergence on undirected, connected sensing graphs. We also note that our treatment of rotational components in the single cycle case is similar in spirit to [17], although the solution in [17] relies primarily on absolute measurements. Our work differs from that presented in [38, 35, 17] in the following ways: (i) we work exclusively with relative measures and thus require no knowledge of an absolute reference frame, (ii) we formulate our cost with respect to a left-invariant metric, allowing us to retain the group structure of $SE(3)$, (iii) we rigorously study the single cycle case and present a closed-form solution which we prove to have certain invariance and optimality properties with respect to both rotation and translation, (iv) we formulate a continuous-time Lyapunov-based algorithm, which can be implemented in a straightforward manner and treats both rotations and translations using similar theoretical machinery, and (v) we prove convergence on any weakly connected digraph and illustrate through simulation that we achieve noise reduction similar to that of an analogous linear problem in the case of "Gaussian-like" noise with small variance.

1.3. Contributions. We choose to work exclusively with relative measurements, since input data in network applications is generally with respect to relative, rather than absolute, reference frames. By doing so, we are forced to explicitly consider cyclic constraints, but we require no knowledge of any global environment; thus, we avoid other potential difficulties (initialization issues, ambiguities, etc.) that may arise due to transformations to a global frame. In this setting, we adopt a dynamical systems perspective on the general problem of sensor network localization by constructing a continuous-time Lyapunov-based gradient flow algorithm. As such, our work provides a novel and valuable contribution to the problem of sensor network localization in the following ways: (i) we provide a unique approach to localizing both positions and orientations in three dimensions, treating both constructs using similar theoretical machinery, (ii) we formulate the problem with respect to only relative measurements in $SE(3)$ and thus require no knowledge of any absolute reference frame, (iii) we rigorously study the single cycle case, providing a closed-form solution and proving invariance and optimality properties with respect to both rotation and translation, (iv) we provide an algorithm for general graphs which is *cycle-distributed* and can be implemented in a straightforward way, providing rigorous convergence proofs which require only weak connectivity of the sensing graph, and (v) we validate our algorithms through comparison with an analogous linear problem.

Specifically, our contributions are as follows. First, we formulate the sensor network localization problem for a single cycle of length n as a constrained least squares optimization over $SE(3)^n$. We present a feasible, closed-form solution which distributes the screw angle error equally among edges in the cycle. Further, we show that when viewed as a parametric function, our solution is intimately related to the minimum length geodesic connecting the composition of displacements in the cycle to the identity. We also show that, in certain cases, namely restriction to pure translation or pure rotation, our solution is invariant to choice of representation of the constraint function. We conclude our theoretical study of the single cycle case by showing that, in some cases, our solution is the unique constrained minimizer of an appropriate cost. Second, we study the localization problem on a general directed and weakly connected sensing graph that contains many semicycles (generalized cycles) and present a solution methodology which is based on the limits of solution curves to a continuous-time ordinary differential equation (ODE). Using Lyapunov-based arguments, we create an algorithm which treats both rotations and translations in a way that is structurally similar in both cases, and we show that solutions obtained using this method satisfy all semicycle constraints. Finally, we present numerical simulations to verify our algorithms and demonstrate their utility. For a single cycle, we compare our closed-form solution to the solution produced by the continuous-time flow for general graphs and show that the two methods produce feasible solutions with nearly identical costs (up to numerical implementation) for the case of pure rotation and pure translation. For general graphs, we demonstrate convergence of our continuous-time solution method for a random geometric graph and show that although our general graph solution is not optimal for general noise models, for "Gaussian-like" noise with small variance, we achieve noise reduction comparable to that achieved by a classic least squares estimator for an analogous linear problem.

1.4. Organization. This paper is organized as follows: section 2 gives background information about relevant topics. Sections 3 and 4 present the localization problem and our proposed solution for a single cycle graph. Sections 5 and 6 present the localization problem and our proposed solution for general sensing graphs. Numerical simulations are in section 7, and proofs of Theorems 3, 4, and 8 are provided in the appendices.

2. Mathematical preliminaries.

In this section, we provide a brief review of matrix Lie groups and graph theory. For more information, the reader is encouraged to refer to [12, 30, 23, 33, 20].

2.1. Matrix Lie groups. Let $\mathbb{0}_n \in \mathbb{R}^n$ denote a vector of zeros and $I_n \in \mathbb{R}^{n\times n}$ denote the identity matrix. Let $SE(3)$ and $SO(3)$ denote the special Euclidean and special orthogonal groups, respectively, and let $\mathfrak{se}(3)$ and $\mathfrak{so}(3)$ denote their respective matrix Lie algebras.

Let $\mathrm{An}(R) \in [0, \pi]$ denote the rotation angle associated to a rotation matrix $R \in SO(3)$. This angle obeys the relation $\mathrm{tr}(R) = 1 + 2\cos(\mathrm{An}(R))$. If $0 < \mathrm{An}(R) < \pi$, let $\mathrm{Ax}(R) \in \mathbb{R}^3$ denote the unique, unit-length rotation axis associated to R. If $\mathrm{An}(R) = 0$, define $\mathrm{Ax}(R) := \mathbb{0}_3$. For $R \in SO(3)$ with $\mathrm{An}(R) < \pi$, we have (i) $R\,\mathrm{Ax}(R) = \mathrm{Ax}(R)$, (ii) $R = \exp(\mathrm{An}(R)\widehat{\mathrm{Ax}(R)})$, where $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ is the isomorphism defined by $\hat{v}w = v \times w$ for any $v, w \in \mathbb{R}^3$, and (iii) $2\widehat{\mathrm{Ax}(R)}\sin(\mathrm{An}(R)) = R - R^T$ (a consequence of Rodrigues' rotation formula). For $R \in SO(3)$ with $\mathrm{An}(R) < \pi$, the *nth root* of R is[1]

$$\sqrt[n]{R} = \exp\left(\frac{1}{n}\,\mathrm{An}(R)\widehat{\mathrm{Ax}(R)}\right) \in SO(3). \tag{1}$$

[1]There are many natural ways to define the nth root of $R \in SO(3)$. Indeed, if $\sqrt[n]{R}$ were defined by replacing the coefficient $1/n$ by $(1 + 2\pi k)/n$ for any $k \in \mathbb{Z}$, the resultant matrix would satisfy $(\sqrt[n]{R})^n = R$. However, the reader should be aware that the results contained herein are specific to the particular definition of the nth root given in the text.

Note that $\mathrm{An}(\sqrt[n]{R}) = \mathrm{An}(R)/n$ and $\mathrm{Ax}(\sqrt[n]{R}) = \mathrm{Ax}(R)$. If $R \in SO(3)$ with $\mathrm{An}(R) = \pi$, there exist exactly two unit-length vectors $\omega_1, \omega_2 \in \mathbb{R}^3$ that satisfy $R = \exp(\pi\widehat{\omega}_1) = \exp(\pi\widehat{\omega}_2)$. These vectors satisfy $\omega_1 = -\omega_2$. For any $R \in SO(3)$, $\omega \in \mathbb{R}^3$, we have the identity $R\widehat{\omega}R^T = \widehat{R\omega}$. Thus for any $R_1, R_2 \in SO(3)$ (including those with rotation angle π), there exists a unit-length $\omega \in \mathbb{R}^3$ such that $R_1R_2R_1^T = \exp(\mathrm{An}(R_2)R_1\widehat{\omega}R_1^T) = \exp(\mathrm{An}(R_2)\widehat{(R_1\omega)})$ and thus $\mathrm{An}(R_1R_2R_1^T) = \mathrm{An}(R_2)$.

For a displacement matrix $E \in SE(3)$, we write $E = (R, p)$ to represent

$$E = \begin{bmatrix} R & p \\ \mathbb{0}_3^T & 1 \end{bmatrix}. \tag{2}$$

If $E = (I_3, p)$ for some $p \in \mathbb{R}^3$, we say that E represents *pure translation*. Similarly, if $E = (R, \mathbb{0}_3)$ for some $R \neq I_3 \in SO(3)$, we say that E represents *pure rotation*. For any $E = (R, p) \in SE(3)$, we have $\mathrm{tr}(E) = 2 + 2\cos(\mathrm{An}(R))$. Note that $\mathrm{tr}(E) = 0$ if and only if $\mathrm{An}(R) = \pi$. In addition to the map already defined for rotations, we use the symbol $\widehat{\cdot}$ to denote the map $\widehat{\cdot} : \mathbb{R}^6 \to \mathfrak{se}(3)$ defined by

$$\widehat{\xi} := \widehat{\begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}} = \begin{bmatrix} \widehat{\xi}_1 & \xi_2 \\ \mathbb{0}_3^T & 0 \end{bmatrix} \in \mathfrak{se}(3),$$

where $\xi_1, \xi_2 \in \mathbb{R}^3$. Define the *screw angle*[2] of $E = (R, p) \in SE(3)$ by

$$\mathrm{An}(E) = \begin{cases} \mathrm{An}(R) & \text{if } R \neq I_3, \\ \|p\| & \text{if } R = I_3. \end{cases} \tag{3}$$

Simple calculation yields $\mathrm{An}(E_1E_2E_1^{-1}) = \mathrm{An}(E_2)$ for any $E_1, E_2 \in SE(3)$. For $E = (R, p)$ with $\mathrm{tr}(E) \neq 0$, define the *screw axis* $\mathrm{Ax}(E) := \mathrm{Ax}(R) \in \mathbb{R}^3$ and the *translational differential* $\mathrm{Tran}(E) \in \mathbb{R}^3$ by

(4)

$$\mathrm{Tran}(E) := \begin{cases} \mathbb{0}_3 & \text{if } E = I_4, \\ p/\|p\| & \text{if } R = I_3, E \neq I_4, \\ \left[(I_3 - R)\widehat{\mathrm{Ax}(E)} + \left(I_3 + \widehat{\mathrm{Ax}(E)}^2\right)\mathrm{An}(E)\right]^{-1} p & \text{otherwise.} \end{cases}$$

The vectors $\mathrm{Tran}(E)$ and $\mathrm{Ax}(E)$ are well defined and satisfy $E = \exp(\mathrm{An}(E)\widehat{\xi})$, where $\xi = [\mathrm{Ax}(E)^T\ \mathrm{Tran}(E)^T]^T$. This representation of the matrix E is unique.

Let $\xi_1 := [\omega_1^T, v_1^T]^T, \xi_2 := [\omega_2^T, v_2^T]^T \in \mathbb{R}^6$. Define the inner product $\langle\cdot,\cdot\rangle_{I_4} : \mathfrak{se}(3) \times \mathfrak{se}(3) \to \mathbb{R}$ by $\langle\widehat{\xi}_1, \widehat{\xi}_2\rangle = \alpha\omega_1^T\omega_2 + \beta v_1^T v_2$, where $\alpha, \beta \in \mathbb{R}_{>0}$. This inner product induces a left-invariant Riemannian metric by associating the tangent space of each $E \in SE(3)$ with the map $\langle\cdot,\cdot\rangle_E : T_ESE(3) \times T_ESE(3) \to \mathbb{R}$, where

$$\left\langle E\widehat{\xi}_1, E\widehat{\xi}_2 \right\rangle_E = \left\langle \widehat{\xi}_1, \widehat{\xi}_2 \right\rangle_{I_4} = \alpha\omega_1^T\omega_2 + \beta v_1^T v_2. \tag{5}$$

With this metric, $SE(3)$ naturally carries the structure of a topological metric space with distance function $d : SE(3) \times SE(3) \to \mathbb{R}_{\geqslant 0}$ defined as the arc length of a minimum length geodesic. Using the explicit definition of the minimum length geodesic

[2]The screw angle $\mathrm{An}(E)$ of a displacement $E \in SE(3)$ does not have a literal interpretation as an angle in the general case, as does the rotation angle $\mathrm{An}(R)$ of a rotation matrix $R \in SO(3)$. However, it is analogous in the following sense: If $R \in SO(3)$ and $E \in SE(3)$, then $R = \exp(\mathrm{An}(R)\widehat{\omega})$ for some unit length $\omega \in \mathbb{R}^3$, and $E = \exp(\mathrm{An}(E)\widehat{\xi})$ for some $\xi := [\xi_1^T, \xi_2^T]^T \in \mathbb{R}^6$, where at least one of $\xi_1, \xi_2 \in \mathbb{R}^3$ is unit length.

with respect to (5) presented in [41], for $E_1 = (R_1, p_1), E_2 = (R_2, p_2)$ we have

$$d(E_1, E_2) = \sqrt{\alpha \operatorname{An}(R_1^T R_2)^2 + \beta \|p_2 - p_1\|^2}. \tag{6}$$

If E_1, E_2 are two displacements both representing pure rotation or both representing pure translation, then $d(E_1, E_2) = c \operatorname{An}(E_1^{-1} E_2)$, where $c = \sqrt{\alpha}$ in the rotation case and $c = \sqrt{\beta}$ in the translation case. The distance function d satisfies the triangle inequality: $d(E_1, E_3) \leqslant d(E_1, E_2) + d(E_2, E_3)$ for any $E_1, E_2, E_3 \in SE(3)$.

Remark 1 (choice of metric). We make use of the left-invariant Riemannian metric (5) and the associated distance function (6) because they are simple, well characterized (e.g., [41]), and intuitive to most applications. Indeed, the distance (6) is simply a weighted sum of the rotational and translational errors. A thorough study of other metrics on $SE(3)$ is beyond the scope of this paper; however, a few comments are in order. First, we could just as easily have chosen the right-invariant Riemannian metric analogous to (5) and generated a set of results virtually identical to those herein. Second, we note that some authors, e.g., [38], choose to discard the group structure of $SE(3)$ and instead work with metrics defined on the product group $SO(3) \times \mathbb{R}^3$. If we take $\alpha = \beta = 1$, then the distance (6) is equivalent to the *double geodesic distance* [5], a commonly used metric defined on $SO(3) \times \mathbb{R}^3$.

2.2. Elements of graph theory. Let $G = (\mathcal{V}, \mathcal{E})$ represent a digraph with vertex set $\mathcal{V} := \{1, \ldots, m\}$ and edge set $\mathcal{E}$. Let $(i, j) \in \mathcal{E}$ denote a directed edge starting from node i and terminating at node j. A *cycle* is an alternating sequence $\langle v_1, (v_1, v_2), v_2, (v_2, v_3), \ldots, (v_r, v_{r+1}), v_{r+1} \rangle$, in which (i) $r \geqslant 3$, (ii) for all $k \in \{1, \ldots, r\}$, $v_k \in \mathcal{V}$ and $(v_k, v_{k+1}) \in \mathcal{E}$, (iii) $v_{r+1} = v_1$, and (iv) all vertices besides v_1, v_{r+1} are pairwise distinct. A *semicycle* on the graph G is a generalized cycle, whose definition arises by replacing condition (ii) with the following: for all $k \in \{1, \ldots, r\}$, $v_k \in \mathcal{V}$, and either (v_k, v_{k+1}) or $(v_{k+1}, v_k) \in \mathcal{E}$. Every cycle on G is a semicycle, but the converse does not hold. If σ is a semicycle on G, an edge $(i, j) \in \mathcal{E}$ is *contained* in σ if either (i, j) or (j, i) appears in the sequence. With slight abuse of notation, we write $(i, j) \in \sigma$ if σ contains (i, j). Define $G_\sigma \subseteq G$ as the *subgraph* whose vertex and edge sets are the sets of vertices and edges that are contained in the semicycle σ. Define an equivalence class by the rule $\sigma \sim \bar{\sigma}$ if the semicycles $\sigma, \bar{\sigma}$ contain the same set of edges. Given a parametrization $\{1, \ldots, n\}$ of $\mathcal{E}$, let $\mathcal{E}_i$ denote the edge with index i. We define the *semicycle vector* of a semicycle σ as the vector $\ell_\sigma \in \{-1, 0, +1\}^n$ for which $(\ell_\sigma)_i = +1$ if $\mathcal{E}_i \in \sigma$ and has an orientation that is consistent with σ, as $(\ell_\sigma)_i = -1$ if edge $\mathcal{E}_i \in \sigma$ and has an orientation that is opposite of σ, and as $(\ell_\sigma)_i = 0$ otherwise.

Given a digraph G, let $G_{\mathrm{UD}} := (\mathcal{V}, \mathcal{E}_{\mathrm{UD}})$ be the undirected graph obtained by ignoring edge orientations. If G is weakly connected, there exists a spanning tree $T_{\mathrm{UD}} \subseteq G_{\mathrm{UD}}$. Let $T \subseteq G$ be the digraph obtained by replacing each edge in T_{UD} by one of its associated edges in $\mathcal{E}$. If T is augmented by a single edge $(i, j) \in \mathcal{E}$, $(i, j) \notin T$, the resulting graph contains at most one equivalence class of semicycles. A set of *fundamental semicycles* is a set of distinct (nonequivalent) semicycles on G which can be formed by the preceding process. Let $\mathcal{L}_f(G)$ denote a maximal set of fundamental semicycles. Any set $\mathcal{L}_f(G)$ is a basis for the semicycle space, and thus for any semicycle σ there exists a semicycle $\bar{\sigma} \sim \sigma$ such that $G_{\bar{\sigma}} = G_{\sigma^1} \triangle G_{\sigma^2} \triangle \cdots \triangle G_{\sigma^n}$, where $\triangle$ denotes the symmetric difference operation, $n \in \mathbb{N}$, and $\sigma^1, \sigma^2, \ldots, \sigma^n \in \mathcal{L}_f(G)$. If G_{UD} is connected, then any $\mathcal{L}_f(G)$ has the property $|\mathcal{L}_f(G)| = |\mathcal{E}_{\mathrm{UD}}| - |\mathcal{V}| + 1$.

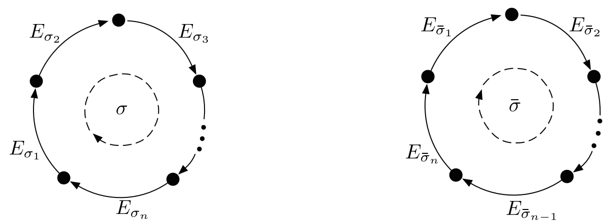

FIG. 2. *The sensing topology for a ring graph $G = (\mathcal{V}, \mathcal{E})$. Note that G contains a single equivalence class of cycles. In the left figure, we represent this equivalence class with the cycle σ, while in the right figure, we represent this equivalence class with a different cycle $\bar{\sigma}$. In this example, we see that $E_{\sigma_1} = E_{\bar{\sigma}_n}$ and $E_{\sigma_i} = E_{\bar{\sigma}_{i-1}}$ for each $i \in \{1, \ldots, n-1\}$. The cycle constraint* (7) *is equivalently represented by the two cycles $\sigma, \bar{\sigma}$.*

3. Problem setup, single cycle case. Consider a network consisting of n sensors, whose sensing capability is represented by a digraph G with edge set $\mathcal{E} = \{(1,2), (2,3), \ldots, (n-1, n), (n, 1)\}$. Parametrize $\mathcal{E}$ by assigning an index $i \in \{1, \ldots, n\}$ to each edge. Let σ denote a cycle on G (note that all cycles on G are equivalent). Let $\sigma_i \in \{1, \ldots, n\}$ be the index of the edge in $\mathcal{E}$ associated with the ith edge in σ.

We assume that each sensor is able to obtain a relative displacement measurement of its neighbor, and thus we associate to each edge $\mathcal{E}_i \in \mathcal{E}$ a displacement matrix $E_i = (R_i, p_i) \in SE(3)$. A composition of true relative displacements around the loop formed by σ must equal the identity matrix. However, we assume that our measurements are corrupted by some noise such that $E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n} \neq I_4$. We aim to find *correction matrices* $C_1, C_2, \ldots, C_n \in SE(3)$ which satisfy the cycle constraint

$$E_{\sigma_1} C_{\sigma_1} E_{\sigma_2} C_{\sigma_2} \cdots E_{\sigma_n} C_{\sigma_n} = I_4 \tag{7}$$

and are optimal in some way. Specifically, we consider the least squares optimization

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{n} d(C_i, I_4)^2 \\ &\text{subject to} && E_{\sigma_1} C_{\sigma_1} E_{\sigma_2} C_{\sigma_2} \cdots E_{\sigma_n} C_{\sigma_n} = I_4, \end{aligned} \tag{8}$$

where $d : SE(3) \times SE(3) \to \mathbb{R}_{\geqslant 0}$ is defined in (6). Note that (7) is the constraint function of the optimization (8). The constraint (7) is satisfied if and only if $(E)^{-1} E_{\sigma_1} C_{\sigma_1} E_{\sigma_2} C_{\sigma_2} \cdots E_{\sigma_n} C_{\sigma_n} (E) = I_4$ for any $E \in SE(3)$. Thus, we can form an alternative, equivalent representation of the constraint (7) by choosing a different cycle $\bar{\sigma}$ with $\bar{\sigma} \sim \sigma$, giving the constraint $E_{\bar{\sigma}_1} C_{\bar{\sigma}_1} E_{\bar{\sigma}_2} C_{\bar{\sigma}_2} \cdots E_{\bar{\sigma}_n} C_{\bar{\sigma}_n} = I_4$ (see Figure 2). We seek to find a solution that is independent of how we choose to represent the cycle constraint.

If $E_1, E_2, \ldots, E_n$ all represent pure rotation, we could formulate a localization problem on $SO(3)^n$ by constraining all correction matrices to represent pure rotation. In this case, the optimization analogous to (8) is

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^{n} \mathrm{An}(C_i)^2 \\ &\text{subject to} && E_{\sigma_1} C_{\sigma_1} E_{\sigma_2} C_{\sigma_2} \cdots E_{\sigma_n} C_{\sigma_n} = I_4, \\ &&& C_i = (R_i, \mathbb{0}_n) \qquad \text{for all } i \in \{1, \ldots, n\}. \end{aligned} \tag{9}$$

Similarly, we can derive an optimization on $\mathbb{R}^3$ if $E_1, \ldots, E_n$ represent pure translation:

$$
\begin{aligned}
&\text{minimize} && \sum_{i=1}^{n} \|p_i\|^2 \\
&\text{subject to} && E_{\sigma_1} C_{\sigma_1} E_{\sigma_2} C_{\sigma_2} \cdots E_{\sigma_n} C_{\sigma_n} = I_4, \\
& && C_i = (I_3, p_i) \qquad \text{for all } i \in \{1, \ldots, n\}.
\end{aligned} \tag{10}
$$

We also seek to study the relationship of our proposed general solutions to the solutions to these specific cases.

Remark 2 (noise model). Our formulation of the optimization (8) is independent of the type of noise present. However, for illustrative purposes (particularly in section 7), we sometimes refer to "Gaussian-like" noise, which we model as follows: assume that the true displacements are given by $\bar{E}_i = (\bar{R}_i, \bar{p}_i) \in SE(3)$ for $i \in \{1, \ldots, n\}$. We assume the nominal measurements are given by

$$E_i = (\bar{R} \exp(\widehat{v_i^r}), \bar{p}_i + v_i^t),$$

where $v_i^r, v_i^t \sim \mathcal{N}(\mathbb{0}_3, \nu^2 I_3)$ are independent and identically distributed Gaussian random variables. For an in-depth discussion of other possible noise models on $SE(3)$, we refer the reader to [10].

4. Solution to the single cycle case. To avoid tedious notation, from this point forward we adopt the following convention for products of indexed matrices.

CONVENTION 1. *Define $E_{\sigma_a} E_{\sigma_{a+1}} \cdots E_{\sigma_b} := I_4$ and $R_{\sigma_a} R_{\sigma_{a+1}} \cdots R_{\sigma_b} := I_3$ whenever* (i) $b < a$, (ii) $b = 0$, *or* (iii) $a > |\sigma|$, *where $|\sigma|$ is the number of edges contained in the cycle or semicycle under consideration. Whenever there is possibility of ambiguity, we group compositions with parentheses. For example, if $|\sigma| = n$, then, with this convention, we have $(E_{\sigma_2} E_{\sigma_3} \cdots E_{\sigma_n})(E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_0}) = E_{\sigma_2} E_{\sigma_3} \cdots E_{\sigma_n}$.*

Specific reasons for using this convention will become clear as the exposition progresses. In short, Convention 1 permits concise, parametric presentations of results containing compositions of displacements, and, therefore, allows us to avoid tedious case-by-case treatments of individual edges based on their associated index.

Consider now the problem setup of section 3. For convenience, we define the set

$$SE(3)^n_{\mathrm{tr}\neq 0} := \{(E_1, E_2, \ldots, E_n) \in SE(3)^n \mid \mathrm{tr}\,(E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n}) \neq 0\}.$$

Adopting Convention 1, we introduce the following definition.

DEFINITION 1 (equal angle solution). *Given $(E_1, E_2, \ldots, E_n) \in SE(3)^n_{\mathrm{tr}\neq 0}$, define the equal angle solution $C^\sigma = (C_1^\sigma, C_2^\sigma, \ldots, C_n^\sigma) \in SE(3)^n$ by*

$$C_i^\sigma = \left(E_{\sigma_{i+1}} E_{\sigma_{i+2}} \cdots E_{\sigma_n}\right) (C_{\mathrm{EA}}^\sigma)_i \left(E_{\sigma_{i+1}} E_{\sigma_{i+2}} \cdots E_{\sigma_n}\right)^{-1},$$

where

$$(C_{\mathrm{EA}}^\sigma)_i = \left(\sqrt[n]{(R_\sigma)^T}\ ,\ -\frac{1}{n} \left(\sqrt[n]{(R_\sigma)^T} \right)^{n-i+1} p_\sigma \right)$$

and $R_\sigma \in SO(3)$, $p_\sigma \in \mathbb{R}^3$ are defined such that $E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n} = (R_\sigma, p_\sigma)$.

We present some key properties of the equal angle solution in Theorem 2.

THEOREM 2 (feasibility and relation to geodesics). *Suppose $(E_1, E_2, \ldots, E_n) \in SE(3)^n_{\mathrm{tr}\neq 0}$. The equal angle solution $C^\sigma \in SE(3)^n$ has the following properties:*

(i) *for each* $i \in \{1, \dots, n\}$, $\mathrm{An}(C_i^\sigma) = \frac{1}{n} \mathrm{An}\left(E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n}\right)$*;*

(ii) $E_{\sigma_1} C_1^\sigma E_{\sigma_2} C_2^\sigma \cdots E_{\sigma_n} C_n^\sigma = I_4$*; and*

(iii) *the curve defined by* $[0,1] \ni t \mapsto E_{\sigma_1} C_1^\sigma(t) E_{\sigma_2} C_2^\sigma(t) \cdots E_{\sigma_n} C_n^\sigma(t)$ *is a minimum length geodesic between* $E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n}$ *and* I_4 *with respect to the metric* (5)*, where* $C_i^\sigma(t)$ *is the parametrized matrix obtained by replacing* $(C_{\mathrm{EA}}^\sigma)_i$ *in Definition* 1 *with* $(C_{\mathrm{EA}}^\sigma)_i(t) := ((R_{\mathrm{EA}}^\sigma)_i(t), (p_{\mathrm{EA}}^\sigma)_i(t))$ *defined by*

$$\begin{aligned}(R_{\mathrm{EA}}^\sigma)_i(t) &= \exp\left(-\frac{1}{n}\mathrm{An}(R_\sigma)\widehat{\mathrm{Ax}(R_\sigma)} t\right),\\ (p_{\mathrm{EA}}^\sigma)_i(t) &= -\frac{1}{n}\exp\left(\left(\frac{(i-1)}{n}t - 1\right)\mathrm{An}(R_\sigma)\widehat{\mathrm{Ax}(R_\sigma)}\right) p_\sigma t.\end{aligned}$$

With this definition, $C_i^\sigma(1) = C_i^\sigma$ *for all* $i \in \{1, \dots, n\}$*.*

Proof. For statement (i), we note that for any $i \in \{1, \dots, n\}$ we have

$$\begin{aligned}\mathrm{An}(C_i^\sigma) &= \mathrm{An}\left((E_{\sigma_{i+1}} E_{\sigma_{i+2}} \cdots E_{\sigma_n})(C_{\mathrm{EA}}^\sigma)_i (E_{\sigma_{i+1}} E_{\sigma_{i+2}} \cdots E_{\sigma_n})^{-1}\right)\\ &= \mathrm{An}((C_{\mathrm{EA}}^\sigma)_i)\\ &= \frac{1}{n}\mathrm{An}\left(E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n}\right).\end{aligned}$$

For statement (ii), we use direct substitution to show

$$\begin{aligned}E_{\sigma_1} C_1^\sigma E_{\sigma_2} C_2^\sigma \cdots E_{\sigma_n} C_n^\sigma &= E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n} (C_{\mathrm{EA}}^\sigma)_1 (C_{\mathrm{EA}}^\sigma)_2 \cdots (C_{\mathrm{EA}}^\sigma)_n\\ &= E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n} \begin{bmatrix} \left(\sqrt[n]{(R_\sigma)^T}\right)^n & -\left(\sqrt[n]{(R_\sigma)^T}\right)^n p \\ \mathbb{0}_3^T & 1 \end{bmatrix}\\ &= E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n} (E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n})^{-1} = I_4.\end{aligned}$$

To prove statement (iii), we use direct substitution once again to show

$$\begin{aligned}&E_{\sigma_1} C_1^\sigma(t) \cdots E_{\sigma_n} C_n^\sigma(t)\\ &= E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n} \begin{bmatrix} \exp\left(-\mathrm{An}(R_\sigma)\widehat{\mathrm{Ax}(R_\sigma)} t\right) & -(R_\sigma)^T p_\sigma t \\ \mathbb{0}_3^T & 1 \end{bmatrix}.\end{aligned}$$

This is a minimum length geodesic between $E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n}$ and I_4 [41]. Setting $t = 1$ and using the definition of $C_i^\sigma(t)$ easily proves the final statement. ∎

Roughly speaking, the solution C^σ attempts to "distribute" a minimum length geodesic between $E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_n}$ and I_4 equally among the edges in the cycle through conjugation. However, since there is no Riemannian metric on $SE(3)$ that is both bi-invariant and positive definite [41], the distance function d is not, in general, invariant under coordinate changes, i.e., conjugation. As a result, in the general case, the equal angle solution may be sensitive to the choice of representation of the constraint function (7) and may not be optimal with respect to the optimization (8). In certain cases, however, we can comment on these two properties, as we show with the following theorems. For clarity, we postpone the proof of these results until Appendix A.

Theorem 3 (invariance). *Suppose* $(E_1, E_2, \dots, E_n) \in SE(3)^n_{\mathrm{tr}\neq 0}$*, and let* $\bar\sigma$ *be a cycle on the graph* G *such that* $\sigma \sim \bar\sigma$*. If* $\mathrm{Ax}((E_{\sigma_k} E_{\sigma_{k+1}} \cdots E_{\sigma_n})(E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_{k-1}}))$ *is aligned*[3] *with* $\mathrm{Tran}((E_{\sigma_k} E_{\sigma_{k+1}} \cdots E_{\sigma_n})(E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_{k-1}}))$ *for all* $k \in \{1, \dots, n\}$*, then* $C_i^\sigma = C_{j(\bar\sigma,\sigma_i)}^{\bar\sigma}$ *for all* $i \in \{1, \dots, n\}$*, where* $j(\bar\sigma, \sigma_i)$ *is the index such that* $\bar\sigma_{j(\bar\sigma,\sigma_i)} = \sigma_i$*.*

[3] We say that $w, v \in \mathbb{R}^3$ are *aligned* if $\widehat{w} v = \mathbb{0}_3$.

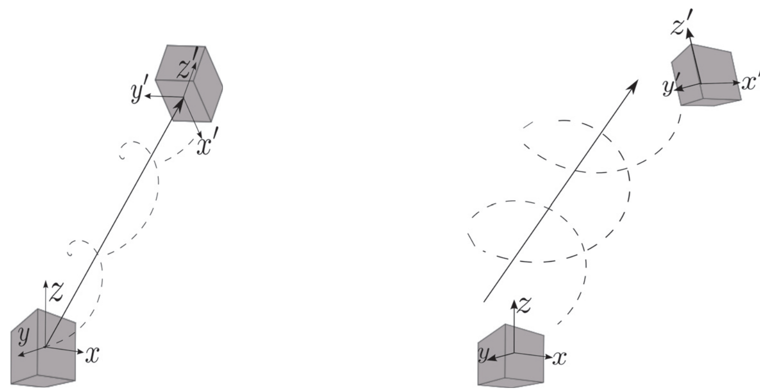

FIG. 3. *Left: An illustration of a screw motion whose rotation axis intersects the origin of the body reference frame. If $E \in SE(3)$ is associated with this screw motion, then* $\mathrm{Tran}(E)$ *is aligned with* $\mathrm{Ax}(E)$*. The case of pure rotation or pure translation is a special case of this phenomena. Right: An illustration of a screw motion for which the rotation axis does not intersect the origin of the body reference frame, in which case* $\mathrm{Tran}(E)$ *is not aligned with* $\mathrm{Ax}(E)$.

THEOREM 4 (optimality). *Suppose* $(E_1, E_2, \ldots, E_n) \in SE(3)^n_{\mathrm{tr}\neq 0}$. *We have*

(i) *if* E_i *represents pure rotation for each* $i \in \{1, \ldots, n\}$*, then the equal angle solution* C^σ *is the unique, global minimizer of the optimization* (9)*, and*

(ii) *if* E_i *represents pure translation for each* $i \in \{1, \ldots, n\}$*, then the equal angle solution* C^σ *is the unique, global minimizer of the optimization* (10).

A few comments are in order. First, any displacement matrix E has an associated rigid body screw motion, and $\mathrm{Ax}(E)$ is aligned with $\mathrm{Tran}(E)$ if and only if the screw motion associated with the displacement E has an axis of rotation that intersects the origin of the body reference frame (see Figure 3) [30, Chapter 2]. Therefore, Theorem 3 applies in the case where $E_1, E_2, \ldots, E_n$ either all represent pure rotation, or all represent pure translation. Second, under the conditions of Theorem 3, the equal angle solution is not sensitive to our choice of representation of the cycle constraint (7). Third, for localization on $SO(3)$ or $\mathbb{R}^3$, the solution C^σ provides the unique, constrained, global minimizer of the appropriate cost function despite nonlinear cost and constraints. Finally, we remark that in the case of localization on $\mathbb{R}^3$, the optimization (10) has a quadratic cost and linear constraints. Therefore, it is possible to use other tools to find a closed-form solution. Theorem 4 guarantees that the equal angle solution is equivalent to any optimal solution derived using other means.

Remark 3 (relation to equation (8)). If $(E_1, E_2, \ldots, E_n) \in SE(3)^n_{\mathrm{tr}\neq 0}$ and E_i represents pure rotation for each $i \in \{1, \ldots, n\}$, then the equal angle solution is also the unique, global minimizer of the more general optimization (8). Indeed, it is easily shown that the minimizer of (8), in this case, is given by a set of correction matrices that each represent pure rotation, and thus optimality of C^σ follows. However, due to the coupling between rotation and translation, as well as the lack of bi-invariance of (5), the same cannot be said for the pure translation case. That is, for the pure translation case, we do not rule out the existence of correction matrices with nontrivial rotational components that produce a lower value of the cost function of (8).

Remark 4 (other costs). The equal angle solution may have optimality properties with respect to cost functions other than that of (8). Although exploring other costs is not our main focus, we note that in the case of localization on $SO(3)^n$ as in (9),

C^σ is also the unique, global minimizer with respect to the optimization

$$\text{(11)} \qquad \begin{array}{ll} \text{minimize} & -\sum_{i=1}^{n} \operatorname{tr}(C_i) \\ \text{subject to} & E_{\sigma_1} C_{\sigma_1} E_{\sigma_2} C_{\sigma_2} \cdots E_{\sigma_n} C_{\sigma_n} = I_4, \\ & C_i = (R_i, \mathbb{0}_n) \qquad \text{for all } i \in \{1,\dots,n\}. \end{array}$$

5. Problem setup for general graphs. Consider a network whose sensing topology is represented by a directed graph $G = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1,\dots,m\}$ and $|\mathcal{E}| = n$, and let $\mathcal{L}_f(G)$ be a maximal set of fundamental semicycles. Once again, parametrize the edge set by assigning an index $i \in \{1,\dots,n\}$ to each edge in $\mathcal{E}$, and associate to each edge $\mathcal{E}_i$ a displacement matrix $E_i = (R_i, p_i) \in SE(3)$. For consistency of measurements, assume that there is at most one directed edge between any two nodes $i, j \in \mathcal{V}$. Recall that the semicycle vector associated with a semicycle σ is denoted ℓ_σ. Given a semicycle σ on G, we define the map $\overline{\cdot}^\sigma : SE(3) \times \{1,\dots,n\} \to SE(3)$ by

$$\text{(12)} \qquad \overline{(E,i)}^\sigma := \begin{cases} E & \text{if } (\ell_\sigma)_i = +1, \\ E^{-1} & \text{if } (\ell_\sigma)_i = -1, \\ I_4 & \text{otherwise.} \end{cases}$$

A composition of true relative displacements around a set of edges forming a semicycle (composed with proper orientation) must equal the identity matrix. We assume that the measurements E_i are corrupted by some noise such that for any semicycle σ, the constraint[4] $\overline{E_{\sigma_1}}^\sigma \overline{E_{\sigma_2}}^\sigma \cdots \overline{E_{\sigma_{|\sigma|}}}^\sigma = I_4$ may not be satisfied. Analogously to the single cycle case, we seek to find correction matrices $C_1, C_2, \dots, C_n$ to compose with the nominal displacements $E_1, E_2, \dots, E_n$, which are optimal in some way and also have the property that the composition of the resulting displacements around *any* semicycle in the graph equals I_4. That is, for any semicycle σ, we wish to satisfy a constraint analogous to the constraint (7), which can be written as

$$\text{(13)} \qquad \overline{E_{\sigma_1} C_{\sigma_1}}^\sigma \overline{E_{\sigma_2} C_{\sigma_2}}^\sigma \cdots \overline{E_{\sigma_{|\sigma|}} C_{\sigma_{|\sigma|}}}^\sigma = I_4,$$

where $|\sigma|$ is the number of edges contained in σ. The following result shows that it is sufficient to satisfy the constraint (13) only over semicycles contained in $\mathcal{L}_f(G)$.

LEMMA 5 (semicycle constraint). *Suppose $C_1, C_2, \dots, C_n \in SE(3)$ are such that for all semicycles $\sigma \in \mathcal{L}_f(G)$, we have $\overline{E_{\sigma_1} C_{\sigma_1}}^\sigma \overline{E_{\sigma_2} C_{\sigma_2}}^\sigma \cdots \overline{E_{\sigma_{|\sigma|}} C_{\sigma_{|\sigma|}}}^\sigma = I_4$. Then, for any other semicycle $\bar\sigma$ defined on the graph G, we have*

$$\overline{E_{\bar\sigma_1} C_{\bar\sigma_1}}^{\bar\sigma} \overline{E_{\bar\sigma_2} C_{\bar\sigma_2}}^{\bar\sigma} \cdots \overline{E_{\bar\sigma_{|\bar\sigma|}} C_{\bar\sigma_{|\bar\sigma|}}}^{\bar\sigma} = I_4.$$

Proof. For the sake of brevity, we only sketch the proof. It is trivial to show that if the semicycle constraint (13) holds for some semicycle on G, then it holds for any other equivalent semicycle. If $\sigma^1, \sigma^2 \in \mathcal{L}_f(G)$, there exist semicycles $\bar\sigma^1 \sim \sigma^1, \bar\sigma^2 \sim \sigma^2$ such that any shared edges between the two are consecutive, i.e., for some $k \in \{0, 1, \dots, n\}$

$$\overline{E_{\bar\sigma^1_1}}^{\bar\sigma^1} = \overline{E_{\bar\sigma^2_1}}^{\bar\sigma^2}, \overline{E_{\bar\sigma^1_2}}^{\bar\sigma^1} = \overline{E_{\bar\sigma^2_2}}^{\bar\sigma^2}, \dots, \overline{E_{\bar\sigma^1_k}}^{\bar\sigma^1} = \overline{E_{\bar\sigma^2_k}}^{\bar\sigma^2},$$

[4]Note that the map $\overline{\cdot}^\sigma$ has an explicit dependence upon an index $i \in \{1,\dots,n\}$. For ease of notation, when the argument $E \in SE(3)$ has an implicit association with an index i, we drop this explicit dependence from our notation; e.g., we write $\overline{E_i}^\sigma = \overline{(E_i, i)}^\sigma$, $\overline{E_i C_i}^\sigma = \overline{(E_i C_i, i)}^\sigma$, etc.

and there exists no $\tilde{k} > k$ such that $\overline{E_{\bar{\sigma}^1_{\tilde{k}}}}^{\bar{\sigma}^1} = \overline{E_{\bar{\sigma}^2_{\tilde{k}}}}^{\bar{\sigma}^2}$. By assumption, we then have

$$\left(\overline{E_{\bar{\sigma}^1_1} C_{\bar{\sigma}^1_1}}^{\bar{\sigma}^1} \cdots \overline{E_{\bar{\sigma}^1_{|\bar{\sigma}^1|}} C_{\bar{\sigma}^1_{|\bar{\sigma}^1|}}}^{\bar{\sigma}^1}\right)^{-1} \left(\overline{E_{\bar{\sigma}^2_1} C_{\bar{\sigma}^2_1}}^{\bar{\sigma}^2} \cdots \overline{E_{\bar{\sigma}^2_{|\bar{\sigma}^2|}} C_{\bar{\sigma}^2_{|\bar{\sigma}^2|}}}^{\bar{\sigma}^2}\right) = I_4.$$

This is sufficient to prove that the semicycle constraint (13) is satisfied for a semicycle σ defined such that $G_\sigma = G_{\sigma^1} \triangle G_{\sigma^2}$, where $\triangle$ denotes the symmetric difference operation. Since $\mathcal{L}_f(G)$ is a basis for the semicycle space, we can assume without loss of generality that the $G_{\bar{\sigma}}$ is equal to the symmetric difference of the subgraphs associated with a finite number of elements of the set $\mathcal{L}_f(G)$. Using logic similar to that above, the statement can be proved by induction. ∎

In light of Lemma 5, we can formally pose the problem for general graphs as a constrained optimization:

$$\text{(14)} \qquad \begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=1}^{n} d(C_i, I_4)^2 \\ \text{subject to} & \overline{E_{\sigma_1} C_{\sigma_1}}^{\sigma} \overline{E_{\sigma_2} C_{\sigma_2}}^{\sigma} \cdots \overline{E_{\sigma_{|\sigma|}} C_{\sigma_{|\sigma|}}}^{\sigma} = I_4 \qquad \text{for all } \sigma \in \mathcal{L}_f(G), \end{array}$$

where d is defined in (6). Note that by Lemma 5, the feasible set is the same regardless of our choice of the set $\mathcal{L}_f(G)$.

6. Solution for general graphs.

In this section, we extend our analysis to a larger class of graphs, and thus we consider the problem setup of section 5.

6.1. Incorporating edge orientations.

Recalling Convention 1, we extend our definition of the equal angle solution as follows.

Definition 6 (equal angle solution (general)). *Given $\sigma \in \mathcal{L}_f(G)$ and nominal measurements $(E_{\sigma_1}, E_{\sigma_2}, \ldots, E_{\sigma_{|\sigma|}})$ satisfying $\operatorname{tr}(\overline{E_{\sigma_1}}^{\sigma} \overline{E_{\sigma_2}}^{\sigma} \cdots \overline{E_{\sigma_{|\sigma|}}}^{\sigma}) \neq 0$, we define $C^\sigma = (C^\sigma_1, C^\sigma_2, \ldots, C^\sigma_{|\sigma|})$ by*

$$C^\sigma_i = \begin{cases} \left(\overline{E_{\sigma_{i+1}}}^{\sigma} \cdots \overline{E_{\sigma_{|\sigma|}}}^{\sigma}\right) (C^\sigma_{\mathrm{EA}})_i \left(\overline{E_{\sigma_{i+1}}}^{\sigma} \cdots \overline{E_{\sigma_{|\sigma|}}}^{\sigma}\right)^{-1} & \text{if } (\ell_\sigma)_{\sigma_i} = +1, \\ \left(\overline{E_{\sigma_i}}^{\sigma} \cdots \overline{E_{\sigma_{|\sigma|}}}^{\sigma}\right) (C^\sigma_{\mathrm{EA}})_i^{-1} \left(\overline{E_{\sigma_i}}^{\sigma} \cdots \overline{E_{\sigma_{|\sigma|}}}^{\sigma}\right)^{-1} & \text{if } (\ell_\sigma)_{\sigma_i} = -1, \end{cases}$$

where

$$(C^\sigma_{\mathrm{EA}})_i = \left(\sqrt[|\sigma|]{(R_\sigma)^T}\ ,\ -\frac{1}{|\sigma|} \left(\sqrt[|\sigma|]{(R_\sigma)^T} \right)^{|\sigma|-i+1} p_\sigma \right)$$

and $\overline{E_{\sigma_1}}^{\sigma} \overline{E_{\sigma_2}}^{\sigma} \cdots \overline{E_{\sigma_{|\sigma|}}}^{\sigma} := (R_\sigma, p_\sigma)$.

Notice that Definition 6 is equivalent to Definition 1 when $\ell_\sigma \in \{+1\}^n$, i.e., when σ is a cycle of length n. Now consider an optimization problem associated to our choice of σ:

$$\text{(15)} \qquad \begin{array}{ll} \text{minimize} & \displaystyle\sum_{i=1}^{|\sigma|} d(C_{\sigma_i}, I_4)^2 \\ \text{subject to} & \overline{E_{\sigma_1} C_{\sigma_1}}^{\sigma} \overline{E_{\sigma_2} C_{\sigma_2}}^{\sigma} \cdots \overline{E_{\sigma_{|\sigma|}} C_{\sigma_{|\sigma|}}}^{\sigma} = I_4. \end{array}$$

Results analogous to those of Theorems 2–4 hold for the generalized solution C^σ with respect to (15). We omit explicit statements and proofs of these results.

6.2. Continuous-time flow for general graphs. In this section, we develop a continuous-time flow whose solution curves determine an appropriate localization for the graph G. For the remainder of this section, we redefine the set

$$SE(3)^n_{\mathrm{tr}\neq 0} := \left\{(E_1, E_2, \ldots, E_n) \mid \mathrm{tr}\left(\overline{E_{\sigma_1}}^{\sigma}\,\overline{E_{\sigma_2}}^{\sigma}\cdots\overline{E_{\sigma_{|\sigma|}}}^{\sigma}\right) \neq 0 \text{ for all } \sigma \in \mathcal{L}_f(G)\right\}.$$

In what follows, we assume that the measurements E_i are functions of a parameter, $t \in \mathbb{R}_{\geqslant 0}$. Consider the following definition.

Definition 7 (equal angle flow). *We define the equal angle flow on the set $SE(3)^n_{\mathrm{tr}\neq 0}$ as the vector field defined by*

$$\begin{cases} \begin{cases} \dot{R}_i &= R_i \displaystyle\sum_{\substack{\sigma\in\mathcal{L}_f(G)\\ \mathcal{E}_i\in\sigma}} \frac{\mathrm{An}(R_\sigma)}{|\sigma|\,(\pi^2 - \mathrm{An}(R_\sigma)^2)} \widehat{\mathrm{Ax}\left(C^{\sigma}_{j(\sigma,i)}\right)} \\ \dot{p}_i &= \mathbb{0}_3 \end{cases} & \textit{if } R_\sigma \neq I_3 \textit{ for some } \sigma \in \mathcal{L}_f(G), \\ \begin{cases} \dot{R}_i &= \mathbb{0}_{3\times 3} \\ \dot{p}_i &= R_i \displaystyle\sum_{\substack{\sigma\in\mathcal{L}_f(G)\\ \mathcal{E}_i\in\sigma}} \frac{\|p_\sigma\|}{|\sigma|}\,\mathrm{Tran}\left(C^{\sigma}_{j(\sigma,i)}\right) \end{cases} & \textit{otherwise,} \end{cases}$$

where $\overline{E_{\sigma_1}}^{\sigma}\,\overline{E_{\sigma_2}}^{\sigma}\cdots\overline{E_{\sigma_{|\sigma|}}}^{\sigma} := (R_\sigma, p_\sigma)$ and $j(\sigma, i) \in \{1, \ldots, |\sigma|\}$ is the index such that $\sigma_{j(\sigma,i)} = i$.

For our purposes, we define a *solution* to the equal angle flow as a continuously differentiable map $(E_1, E_2, \ldots, E_n) : \mathbb{R}_{\geqslant 0} \to SE(3)^n_{\mathrm{tr}\neq 0}$, whose domain can be extended to an open set $\mathbb{R}_{>-\delta}$ for some $\delta > 0$ in such a way that the extended map obeys the dynamics defined by the equal angle flow for all $t \in \mathbb{R}_{\geqslant 0}$. We now characterize the well-posedness and convergence of solutions to the equal angle flow. We once again postpone proof of this result until Appendix B.

Theorem 8 (solutions to the equal angle flow). *Assume that $G = (\mathcal{V}, \mathcal{E})$ is weakly connected. Given initial measurements $(E_1(0), E_2(0), \ldots, E_n(0)) \in SE(3)^n_{\mathrm{tr}\neq 0}$, there exists a unique solution $[0, \infty) \ni t \mapsto (E_1(t), E_2(t), \ldots, E_n(t))$ to the equal angle flow. The solution has the properties*

(i) $E_i^* := \lim_{t\to\infty} E_i(t)$ *exists for each $i \in \{1, \ldots, n\}$, and*
(ii) $\overline{E^*_{\sigma_1}}^{\sigma}\,\overline{E^*_{\sigma_2}}^{\sigma}\cdots\overline{E^*_{\sigma_{|\sigma|}}}^{\sigma}$ *represents pure translation for all $\sigma \in \mathcal{L}_f(G)$.*

If, in addition, the measurements $E_1(0), E_2(0), \ldots, E_n(0)$ are such that the composition $\overline{E_{\sigma_1}(0)}^{\sigma}\,\overline{E_{\sigma_2}(0)}^{\sigma}\cdots\overline{E_{\sigma_{|\sigma|}}(0)}^{\sigma}$ represents pure translation for all $\sigma \in \mathcal{L}_f(G)$, then

(iii) $\overline{E^*_{\sigma_1}}^{\sigma}\,\overline{E^*_{\sigma_2}}^{\sigma}\cdots\overline{E^*_{\sigma_{|\sigma|}}}^{\sigma} = I_4$ *for all $\sigma \in \mathcal{L}_f(G)$.*

Theorem 8 motivates a natural method for finding feasible configurations given nominal displacement measurements on any weakly connected digraph G. Indeed, allowing the displacements to evolve according to the equal angle flow guarantees convergence to a set of displacements whose compositions around the semicycles represent pure translation. If we then initialize a second instance of the equal angle flow, using the limiting displacements from the first instance as initial points, we will produce convergence to a set of displacements satisfying all semicycle constraints. We refer to this method of finding a solution as the *equal angle flow method*.

If a semicycle $\sigma \in \mathcal{L}_f(G)$ is disjoint from the rest of the semicycles in $\mathcal{L}_f(G)$, we can compare the configuration produced by the equal angle flow method to the equal angle solution for a single cycle (Definition 6). If, for such a σ, we have $E_{\sigma_i} = (R_{\sigma_i}, \mathbb{0}_3)$

for all $i \in \{1, \ldots, |\sigma|\}$, then under the assumptions of Theorem 8, $E^*_{\sigma_i} = E_{\sigma_i} C^\sigma_i$ for each $i \in \{1, \ldots, |\sigma|\}$, where $E^*_{\sigma_i}$ is the displacement resulting from the equal angle flow method. This is because the flow causes the solution curve to move along the minimum length geodesic with respect to pure rotation. A similar result holds in the case where each $E_{\sigma_i} = (I_3, p_{\sigma_i})$. However, in the general case, we cannot guarantee that the result of the equal angle flow method and the equal angle solution will be equal, since the rotational and translational components of compositions around semicycles may be "coupled"; i.e., altering the rotational component may affect the translational component. We conclude this section with some remarks.

Remark 5 (weighting). Each summand in Definition 7 is weighted by the inverse of its associated semicycle length. These weights can be replaced by any strictly positive set of values, and the convergence properties of Theorem 8 still hold. The reason for our particular choice of weights is twofold: first, our choice of weights helps to enforce consistency between the definitions of the equal angle flow and the solution to the single cycle case, the equal angle solution. Second, in light of Theorem 8, our choice of weights helps to ensure that the errors due to each semicycle are treated equally; that is, in general graphs our weights serve to balance the rate of error convergence across semicycles. To see this, we note that if all semicycles were weighted equally and the equal angle flow method were used to find a solution of two single loop graphs individually, each with the same error, then the graph with more edges would converge faster.

Remark 6 (choice of semicycle basis). If we replace the set $\mathcal{L}_f(G)$ in Definition 7 with any other semicycle basis $\mathcal{B}$ whose elements each contain at least one edge that is not contained in any other element of $\mathcal{B}$ ($\mathcal{L}_f(G)$ always has this property), then the results of Theorem 8 still hold. That is, replacing $\mathcal{L}_f(G)$ with $\mathcal{B}$, we can find feasible configurations using a method analogous to the equal angle flow method.

Note, however, that limiting configurations produced in this way will, in the general case, be affected by the particular choice of $\mathcal{B}$. Indeed, it is straightforward to construct examples in which two different choices of bases cause convergence to two different configurations (e.g., consider a pure rotational planar graph whose semicycle bases contain two semicycles). This is in contrast to the special case of a sensing graph containing a single semicycle, in which choice of basis will not affect the final result of the equal angle flow due to the invariance property of Theorem 3.

Remark 7 (alternative algorithms and distributed properties). If the composition of nominal displacements around each semicycle represents pure translation and we restrict our attention to purely translational corrections, then the optimization (14) can be reformulated as a convex problem. This problem can be solved using standard optimization or linear algebra tools. Thus, an alternative solution approach for general localization is to let displacements follow the equal angle flow until a configuration is reached in which the rotational constraint around all semicycles is satisfied, and subsequently solving the remaining translational optimization using alternative means. Due to dependencies of translational components on rotational components, however, it may be tedious to properly formulate such a standard translational optimization, especially in the case of large graphs with many semicycles. In such cases, the equal angle flow method may offer more straightforward implementation, since it is *semicycle distributed*. That is, the flow at a given node only requires knowledge of the displacements of other nodes with which it shares a cycle. In general, tractability of any method with respect to a specific problem will need to be assessed by system designers.

7. Simulations. In this section, we present a series of simulation studies to illustrate our findings and the behavior of our solutions. For the simulations, it is necessary to choose appropriate α, β values in the definition of the distance function (6). When choosing these values, the larger the ratio α/β, the more penalty is assigned to errors in rotation. Since the rotational and translational components are measured in different units, choosing α, β will generally be application specific. However, we note that, in general, the larger the average translational component present in the graph, the larger the ratio α/β that is necessary to achieve significant noise reduction using our solution methods. Indeed, for larger translational distances, errors in rotation often have a more significant impact on the configuration error than in the case where these distances are small. For our simulations, we illustrate our results for $\alpha = 10, \beta = 1$, and average translational components between 5 and 10 units.

In order to simulate the continuous-time ODE in Definition 7, we discretize the solution by making use of an Euler step. For an ODE defined on $SO(3)$ of the form $\dot{R} = R\widehat{\Omega}$, the Euler step takes the form $R(k+1) = R(k)\exp(\epsilon\widehat{\Omega})$ for some $\epsilon > 0$. Practical implementation of the equal angle flow method involves setting an error threshold for convergence, and once the solution comes sufficiently close to satisfying the constraints of the problem, we use the value of the displacements as the limit.

7.1. Single cycle case. In our first simulation study, we consider the case of a "ring" graph as in section 3. We consider three types of displacement measurements: pure rotation, pure translation, and general displacement. In each simulation, we initialize the noisy displacement measurements so that they do not satisfy the cycle constraint. We then find both the equal angle solution, as well as the solution resulting from the equal angle flow method. Figure 4 shows the cost and constraint function evaluated for both solution methods for varying numbers of nodes. In all cases, both solutions are feasible, i.e., they satisfy the cycle constraint. Further, in the case of pure translation and pure rotation, the equal angle flow method produces costs which are very close to those produced by the equal angle solution, with only small errors due to numerical implementation. In the case of general displacements, the two solutions differ, which is a result of the fact that the equal angle flow method adjusts rotational and translational components separately.

7.2. General graph case. In our second simulation study, we fix a random digraph with 10 nodes. We associate to each edge a displacement matrix and use the equal angle flow method to find new displacements. The left plot in Figure 5 illustrates a typical evolution of the constraint $\sum_\sigma \|I_4 - \overline{E_{\sigma_1}}^\sigma(k) \cdots \overline{E_{\sigma_{|\sigma|}}}^\sigma(k)\|$ (note that requiring this expression to equal 0 is equivalent to enforcing the constraint in (14)), as well as a typical evolution of the cost function of (14) evaluated by letting $C_i(k) = (E_i(0))^{-1}E_i(k)$ at each discrete time step. The plot clearly shows the transition from the first implementation of the equal angle flow to the second implementation, which occurs at $k \approx 1200$. Note that the cost increases monotonically, while the constraint decreases monotonically only during the second implementation of the flow. Since the translational components do not change during the first implementation of the flow, the constraint need not decrease during this phase.

It is clear that for general noise models our solutions for the localization problem on $SE(3)$ are not proven to be optimal. However, for measurements that are subject to "Gaussian-like" noise models with zero mean and small standard deviation $\nu \ll \pi$, and when $\alpha \gg \beta$ in (6), we believe that our presented solution method (i) substantially reduces the mean square error for graphs with large numbers of semicycles, and (ii) in

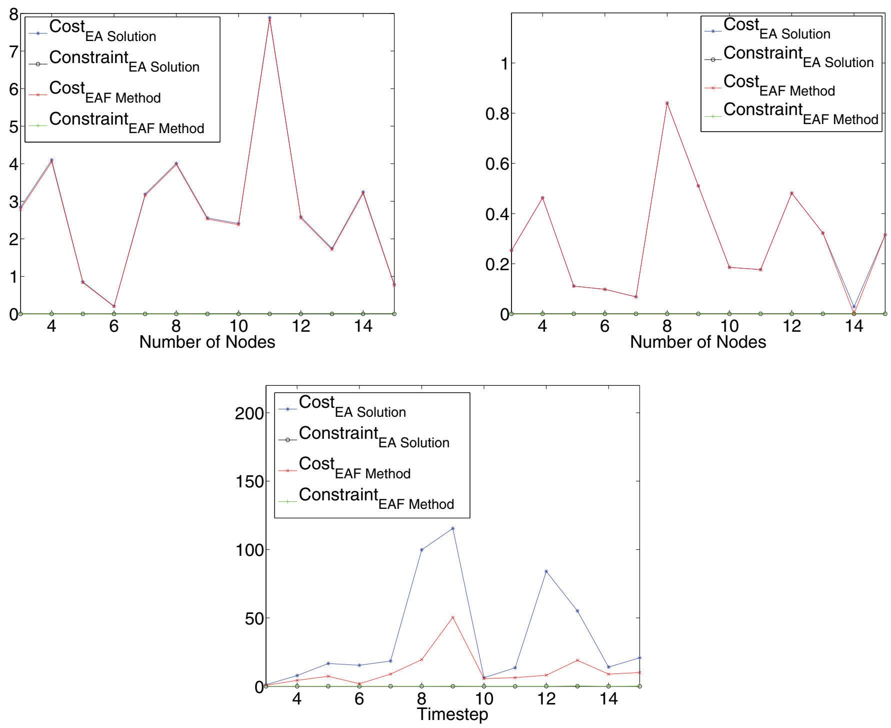

FIG. 4. *We consider the pure rotation case (top left), the pure translation case (top right), and the general displacement case (bottom). For each case, we show the cost and the constraint functions evaluated for the equal angle solution and the solution resulting from the equal angle flow method. The constraint is computed as* $\|E_{\sigma_1}C_{\sigma_1}\cdots E_{\sigma_n}C_{\sigma_n} - I_4\|$.

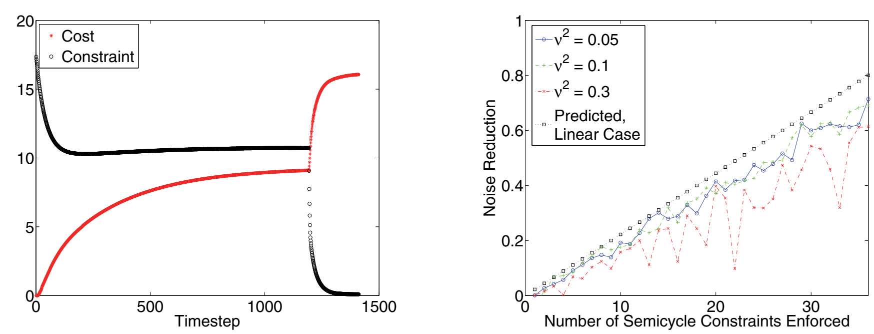

FIG. 5. *Left: The evolution of the constraint* $\sum_\sigma \|I_4 - \overline{E_{\sigma_1}}^\sigma(k)\cdots\overline{E_{\sigma_{|\sigma|}}}^\sigma(k)\|$*, as well as the evolution of the cost function of* (14) *evaluated by letting* $C_i(k) = (E_i(0))^{-1}E_i(k)$ *at each discrete time step* k *under the equal angle flow for a random, weakly connected graph with* 10 *nodes. We assume the noise model described in Remark* 2*. Note the transition from the first implementation of the equal angle flow to the second implementation, which occurs at* $k \approx 1200$*, and note that the trajectory converges to a feasible point as* $k \to \infty$*. Right: An illustration of* 1 *minus the ratio of the mean square error of the configuration resulting from the equal angle flow method to the value of the mean square error of the initial noisy measurements as the number of semicycle constraints enforced increases. For this study, we consider a complete graph with* $m = 10$ *nodes. The plot shows results for* $\nu^2 = 0.05, 0.1, 0.3$ *(see Remark* 2*), as well as the expected noise reduction for the analogous linear problem as predicted by* (17).

terms of noise reduction, is comparable to the classic least squares estimator for an analogous linear problem. We make these statements precise in what follows. The following analysis is analogous to that appearing in [33].

We first construct the linear analogue of problem (14). Since $SE(3)^n$ is a $6n$-dimensional manifold, we consider a problem in which we have noisy measurements $y = p + vI_{6n}$, where $p \in \mathbb{R}^{6n}$ is the nominal measurement vector and $v \in \mathbb{R}$ is a zero mean Gaussian random variable with variance ν^2. We consider additive semicycle constraints, and thus we have a problem of the form

$$\begin{aligned} &\text{minimize} && \|y - \hat{p}\|^2 \\ &\text{subject to} && \mathbf{1}_{\ell_\sigma}^T \hat{p} = 2\pi \mathbf{k}_\sigma \quad \text{for all } \sigma \in \mathcal{L}_f(G), \end{aligned} \tag{16}$$

where $\mathbf{k}_\sigma \in \mathbb{N}^6$ and $\mathbf{1}_{\ell_\sigma} \in \mathbb{R}^{6n\times 6}$ is a vector that is formed by replacing each $+1$ entry in ℓ_σ with I_6, each -1 entry in ℓ_σ with $-I_6$, and each 0 entry with $\mathbb{0}_{6\times 6}$. Suppose p is the true solution. Let $C_f \in \mathbb{R}^{6r\times 6n}$ be the matrix formed by stacking the matrices $\mathbf{1}_{\ell_\sigma}^T$, where $r = |\mathcal{L}_f(G)|$. The true solution p satisfies $p = p_0 + U\eta$, where $C_f p_0 = 2\pi\mathbf{k}_f$, $k_f \in \mathbb{N}^{6r}$ is formed by stacking the vectors $\mathbf{k}_\sigma$, $U \in \mathbb{R}^{6n\times(6n-6r)}$ spans the null space of C_f, and $\eta \in \mathbb{R}^{6n-6r}$ is computed accordingly. It follows that the constraint in (16) can be written as $C_f\hat{p} = 2\pi\mathbf{k}_f$. We calculate the error covariance matrix $Q \in \mathbb{R}^{6n\times 6n}$:

$$Q = \mathbb{E}\left[(\hat{p} - p)(\hat{p} - p)^T\right] = \nu^2 U(U^TU)^{-1}U^T.$$

The matrix $U(U^TU)^{-1}U^T$ is an orthogonal projector onto the vector space that is spanned by the columns of U. The trace of a projector is equal to the rank of the projection, that is, $\operatorname{tr}(U(U^TU)U^T) = 6n - 6r$, and the expected value of the estimation mean square error (MSE) is given by

$$\text{MSE} = \frac{\operatorname{tr}(Q)}{(\text{number of measurements})} = \nu^2 \frac{n-r}{n}. \tag{17}$$

Since r is the number of independent semicycles, for a connected graph we have $r = n - m + 1$, with n, m the number of edges and nodes, respectively. Thus $\text{MSE} = \nu^2 \frac{m-1}{n}$. For a complete graph, $n = m(m-1)/2$, and thus $\text{MSE} = \nu^2 \frac{2}{m}$, which is a noise reduction of at least 80% when $m \geqslant 10$ nodes. Note that once the number of nodes and edges, m, n, along with the noise covariance ν^2 are fixed, as the number of semicycle constraints that we enforce increases, the redundancy of the available noisy measurements is correctly exploited, and we get increased noise reduction.

The right-hand plot in Figure 5 illustrates that our nonlinear problem (14) exhibits similar behavior to that of its linear analogue. Here we consider a complete graph with $m = 10$ nodes and assume that ν is fixed (see Remark 2). We plot 1 minus the ratio of the mean square error of our estimate, which was found using the equal angle flow method, to the value of the mean square error of the initial measurements as we increase the number of semicycle constraints that we choose to enforce. We see that the noise reduction obtained by the equal angle flow method is similar to that of the linear analogue for small ν.

We note that we do not achieve noise reduction comparable to the linear case when α/β is small compared to the average translational component length of the graph. In such cases, other methods, e.g., similar to those presented in [38, 35], may be more appropriate. However, such methods bring about a different set of computational challenges, and thus tractability will need to be assessed by system designers.

8. Conclusion. We have presented a rigorous discussion of the sensor network localization problem on $SE(3)^n$. We started by generating a closed form solution for the case of a graph containing a single cycle of length n. We showed that our solution satisfies the cycle constraint induced by the network sensing topology and distributes the screw angle error equally among the edges. Moreover, in certain cases, namely pure rotation and pure translation, our solution does not depend on the particular representation of the cycle constraint function and is the unique minimizer of an appropriate cost. We then considered the case of a general, weakly connected graph. We showed that our single cycle solution can be adapted to take into account arbitrary edge orientations on any given semicycle, and used this result to generate a continuous-time ODE, for which we can use solution curves to find a feasible localization on the entire graph. We presented a series of simulation studies to verify our results, including a study which shows that for certain noise models, our solution obtains noise reduction comparable to that of the classic least squares for the analogous linear problem.

A variety of future problems are of interest. One possible problem is the issue of implementing these procedures in conjunction with standard robotic operations, such as camera network calibration. In addition, it would also be of interest to explore other cost functions and see how these alternative formulations affect aspects like noise reduction. It would also be useful to find other statistical measures of uncertainty in $SE(3)$ and incorporate these measures into our algorithms. Finally, deriving a flow that alters both the rotational and translational components simultaneously and is guaranteed to converge for general graphs is an open problem that could potentially improve performance.

Appendix A. Proof of Theorems 3 and 4. For this appendix, assume the problem setup and notation of sections 3 and 4. In addition, we introduce the following shorthand: $C_i^\sigma \coloneqq (R_i^\sigma, p_i^\sigma)$, and

$$\mathbf{E}_{i,j}^\sigma \coloneqq (\mathbf{R}_{i,j}^\sigma, \mathbf{p}_{i,j}^\sigma) \coloneqq \begin{cases} (E_{\sigma_i} E_{\sigma_{i+1}} \cdots E_{\sigma_n})(E_{\sigma_1} E_{\sigma_2} \cdots E_{\sigma_j}) & \text{if } j < n, \\ E_{\sigma_i} E_{\sigma_{i+1}} \cdots E_{\sigma_n} & \text{otherwise,} \end{cases}$$

where $i \in \{1, \ldots, n+1\}$ and $j \in \{0, 1, \ldots, n\}$ (recall Convention 1 for products of indexed matrices). Direct calculation yields

$$\mathbf{R}_{i,j}^\sigma = \begin{cases} (R_{\sigma_i} R_{\sigma_{i+1}} \cdots R_{\sigma_n})(R_{\sigma_1} R_{\sigma_2} \cdots R_{\sigma_j}) & \text{if } j < n \\ R_{\sigma_i} R_{\sigma_{i+1}} \cdots R_{\sigma_n} & \text{otherwise,} \end{cases} \tag{18}$$

$$\mathbf{p}_{i,j}^\sigma = \begin{cases} \displaystyle\sum_{k=i}^{n} R_{\sigma_i} R_{\sigma_{i+1}} \cdots R_{\sigma_{k-1}} p_{\sigma_k} + \sum_{k=1}^{j} \mathbf{R}_{i,k-1}^\sigma p_{\sigma_k} & \text{if } j < n, \\ \displaystyle\sum_{k=i}^{n} R_{\sigma_i} R_{\sigma_{i+1}} \cdots R_{\sigma_{k-1}} p_{\sigma_k} & \text{otherwise,} \end{cases} \tag{19}$$

where empty summations in (19) are understood to equal zero, i.e., $\sum_a^b(\cdot) \coloneqq \mathbb{0}_3$ whenever $b < a$. We will need the following lemma.

LEMMA 9 (conjugation of tangent vectors). *If* $(E_1, \ldots, E_n) \in SE(3)^n_{\mathrm{tr}\neq 0}$, *then*

$$E_{\sigma_1} \widehat{\begin{bmatrix} \mathrm{Ax}(\mathbf{E}_{2,1}^\sigma) \\ \mathrm{Tran}(\mathbf{E}_{2,1}^\sigma) \end{bmatrix}} (E_{\sigma_1})^{-1} = \widehat{\begin{bmatrix} \mathrm{Ax}(\mathbf{E}_{1,n}^\sigma) \\ \mathrm{Tran}(\mathbf{E}_{1,n}^\sigma) \end{bmatrix}}.$$

Proof. Simple algebra yields

$$
\begin{aligned}
&E_{\sigma_1}\widehat{\begin{bmatrix}\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})\\ \mathrm{Tran}(\mathbf{E}^{\sigma}_{2,1})\end{bmatrix}}(E_{\sigma_1})^{-1}\\
&\quad=\begin{bmatrix} R_{\sigma_1}\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})}R^T_{\sigma_1} & R_{\sigma_1}\,\mathrm{Tran}(\mathbf{E}^{\sigma}_{2,1})-R_{\sigma_1}\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})}R^T_{\sigma_1}p_{\sigma_1}\\ \mathbb{0}_3^T & 0\end{bmatrix}.
\end{aligned}\tag{20}
$$

Recall that $\mathrm{An}(\mathbf{R}^{\sigma}_{2,1})=\mathrm{An}(\mathbf{R}^{\sigma}_{1,n})$ and $\mathrm{An}(\mathbf{E}^{\sigma}_{2,1})=\mathrm{An}(\mathbf{E}^{\sigma}_{1,n})$, since both rotation and screw angles are invariant under conjugation (see section 2.1). There are two possible cases.

Case I *(*$\mathrm{An}(\boldsymbol{R}^{\sigma}_{2,1})=\mathrm{An}(\boldsymbol{R}^{\sigma}_{1,n})=0$*).* In this case, $\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})=\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})=\mathbb{0}_3$. If $\mathbf{E}^{\sigma}_{2,1}=I_4$, then $\mathrm{Tran}(\mathbf{E}^{\sigma}_{2,1})=\mathrm{Tran}(\mathbf{E}^{\sigma}_{1,n})=\mathbb{0}_3$, and the statement holds trivially. If $\mathbf{E}^{\sigma}_{2,1}\neq I_4$, from (4) and (19) we have

$$
R_{\sigma_1}\,\mathrm{Tran}(\mathbf{E}^{\sigma}_{2,1})=R_{\sigma_1}\frac{\mathbf{p}^{\sigma}_{2,1}}{\mathrm{An}(\mathbf{E}^{\sigma}_{2,1})}=\frac{1}{\mathrm{An}(\mathbf{E}^{\sigma}_{1,n})}\left(\mathbf{R}^{\sigma}_{1,n}p_{\sigma_1}+\sum_{k=2}^{n}\mathbf{R}^{\sigma}_{1,k-1}p_{\sigma_k}\right)\tag{21}
$$

$$
=\frac{1}{\mathrm{An}(\mathbf{E}^{\sigma}_{1,n})}\sum_{k=1}^{n}\mathbf{R}^{\sigma}_{1,k-1}p_{\sigma_k}\tag{22}
$$

$$
=\frac{\mathbf{p}^{\sigma}_{1,n}}{\mathrm{An}(\mathbf{E}^{\sigma}_{1,n})}=\mathrm{Tran}(\mathbf{E}^{\sigma}_{1,n}),\tag{23}
$$

where we have used (i) the invariance of screw angles under conjugation, and (ii) the fact that $\mathrm{An}(\mathbf{R}^{\sigma}_{1,n})=0$ and thus $\mathbf{R}^{\sigma}_{1,n}=I_3$ by assumption. From (20), it follows that

$$
E_{\sigma_1}\widehat{\begin{bmatrix}\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})\\ \mathrm{Tran}(\mathbf{E}^{\sigma}_{2,1})\end{bmatrix}}(E_{\sigma_1})^{-1}=\begin{bmatrix}\widehat{\mathbb{0}}_3 & R_{\sigma_1}\,\mathrm{Tran}(\mathbf{E}^{\sigma}_{2,1})\\ \mathbb{0}_3^T & 0\end{bmatrix}=\widehat{\begin{bmatrix}\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})\\ \mathrm{Tran}(\mathbf{E}^{\sigma}_{1,n})\end{bmatrix}}.
$$

Case II *(*$\mathrm{An}(\boldsymbol{R}^{\sigma}_{2,1})=\mathrm{An}(\boldsymbol{R}^{\sigma}_{1,n})\neq 0$*).* In this case, we examine the entries of the matrix in (20) separately. Regarding the top left entry of the matrix in (20), Rodrigues' rotation formula (see section 2.1) implies

$$
R_{\sigma_1}\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})}R^T_{\sigma_1}=R_{\sigma_1}\frac{\mathbf{R}^{\sigma}_{2,1}-(\mathbf{R}^{\sigma}_{2,1})^T}{2\sin(\mathrm{An}(\mathbf{R}^{\sigma}_{2,1}))}R^T_{\sigma_1}=\frac{\mathbf{R}^{\sigma}_{1,n}-(\mathbf{R}^{\sigma}_{1,n})^T}{2\sin(\mathrm{An}(\mathbf{R}^{\sigma}_{1,n}))}=\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})},
$$

where we once again have used the invariance of rotation angles under conjugation. It remains to show that the top right term of the matrix in (20) is equal to $\mathrm{Tran}(\mathbf{E}^{\sigma}_{1,n})$. Recall (4), and let $\boldsymbol{\xi}:=(I_3-\mathbf{R}^{\sigma}_{1,n})\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}+(I_3+\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}^2)\,\mathrm{An}(\mathbf{E}^{\sigma}_{1,n})$. We have

$$
\begin{aligned}
&R_{\sigma_1}\,\mathrm{Tran}(\mathbf{E}^{\sigma}_{2,1})-R_{\sigma_1}\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})}R^T_{\sigma_1}p_{\sigma_1}\\
&=R_{\sigma_1}[(I_3-\mathbf{R}^{\sigma}_{2,1})\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})}+(I_3+\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{2,1})}^2)\,\mathrm{An}(\mathbf{E}^{\sigma}_{2,1})]^{-1}\mathbf{p}^{\sigma}_{2,1}-\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1}\\
&=R_{\sigma_1}[R^T_{\sigma_1}(I_3-\mathbf{R}^{\sigma}_{1,n})\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}R_{\sigma_1}+R^T_{\sigma_1}(I_3+\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}^2)R_{\sigma_1}\,\mathrm{An}(\mathbf{E}^{\sigma}_{1,n})]^{-1}\mathbf{p}^{\sigma}_{2,1}\\
&\qquad\qquad-\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1}\\
&=R_{\sigma_1}\left[I_3+\left(R^T_{\sigma_1}\boldsymbol{\xi}-R^T_{\sigma_1}\right)R_{\sigma_1}\right]^{-1}\mathbf{p}^{\sigma}_{2,1}-\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1}\\
&=\left[I_3+R_{\sigma_1}\left(R^T_{\sigma_1}\boldsymbol{\xi}-R^T_{\sigma_1}\right)\right]^{-1}R_{\sigma_1}\mathbf{p}^{\sigma}_{2,1}-\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1}\\
&=\boldsymbol{\xi}^{-1}R_{\sigma_1}\left(\mathbf{R}^{\sigma}_{2,0}p_{\sigma_1}+\sum_{k=2}^{n}R_{\sigma_2}R_{\sigma_3}\cdots R_{\sigma_{k-1}}p_{\sigma_k}\right)-\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1}
\end{aligned}
$$

$$= \boldsymbol{\xi}^{-1}\left(\mathbf{R}^{\sigma}_{1,n}p_{\sigma_1} + \sum_{k=2}^{n} R_{\sigma_1}R_{\sigma_2}\cdots R_{\sigma_{k-1}}p_{\sigma_k}\right) - \widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1}$$
$$= \boldsymbol{\xi}^{-1}(\mathbf{R}^{\sigma}_{1,n}p_{\sigma_1} + \mathbf{p}^{\sigma}_{1,n} - p_{\sigma_1}) - \widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1}$$
$$= \mathrm{Tran}(\mathbf{E}^{\sigma}_{1,n}) + \boldsymbol{\xi}^{-1}(\mathbf{R}^{\sigma}_{1,n} - I_3)p_{\sigma_1} - \widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1},$$

where the fourth equality follows from the fact that $P(I_3 + QP)^{-1} = (I_3 + PQ)^{-1}P$ for any $P, Q \in \mathbb{R}^{3\times 3}$ such that $(I_3 + PQ), (I_3 + QP)$ are invertible. Invertibility of the appropriate matrices follows from well-posedness of (4) [30]. To conclude, we show that $\boldsymbol{\xi}^{-1}(\mathbf{R}^{\sigma}_{1,n} - I_3)p_{\sigma_1} - \widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1} = \mathbb{0}_3$. We utilize the following identities [30, Lemma 2.3]: for any $w \in \mathbb{R}^3$, we have $\widehat{w}^2 = ww^T - \|w\|I_3$ and $\widehat{w}^3 = -\|w\|^2\widehat{w}$. By assumption, $\mathrm{An}(\mathbf{R}^{\sigma}_{1,n}) \neq 0$, and thus $\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})$ has unit norm. We have

$$(\mathbf{R}^{\sigma}_{1,n} - I_3)p_{\sigma_1} - \boldsymbol{\xi}\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}p_{\sigma_1}$$
$$= (\mathbf{R}^{\sigma}_{1,n} - I_3)p_{\sigma_1} - [(I_3 - \mathbf{R}^{\sigma}_{1,n})\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}^2 + (\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})} + \widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})}^3)\,\mathrm{An}(\mathbf{E}^{\sigma}_{1,n})]p_{\sigma_1}$$
$$= (\mathbf{R}^{\sigma}_{1,n} - I_3)p_{\sigma_1} - [(I_3 - \mathbf{R}^{\sigma}_{1,n})(\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})\,\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})^T - I_3)]p_{\sigma_1}$$
$$= (\mathbf{R}^{\sigma}_{1,n} - I_3)p_{\sigma_1} - (\mathbf{R}^{\sigma}_{1,n} - I_3)p_{\sigma_1} = \mathbb{0}_3,$$

where the last equality follows from the identity $\mathbf{R}^{\sigma}_{1,n}\,\mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n}) = \mathrm{Ax}(\mathbf{E}^{\sigma}_{1,n})$. ∎

Proof of Theorem 3. For any cycle $\bar{\sigma}$ on the graph G with $\bar{\sigma} \sim \sigma$, either $\mathbf{E}^{\sigma}_{2,1} = \mathbf{E}^{\bar{\sigma}}_{1,n}$, or there exist $k \in \{1,\ldots,n-2\}$ and a sequence of cycles $\sigma^1, \sigma^2, \ldots, \sigma^k$ such that

$$\sigma \sim \sigma^1 \sim \sigma^2 \sim \cdots \sim \sigma^k \sim \bar{\sigma}, \tag{24}$$

$\mathbf{E}^{\sigma}_{2,1} = \mathbf{E}^{\sigma^1}_{1,n}$, $\mathbf{E}^{\sigma^k}_{2,1} = \mathbf{E}^{\bar{\sigma}}_{1,n}$, and $\mathbf{E}^{\sigma^j}_{2,1} = \mathbf{E}^{\sigma^{j+1}}_{1,n}$ for all $j \in \{1,\ldots,k-1\}$. Thus, it suffices to show that for any $i \in \{1,\ldots,n\}$, $C^{\sigma}_i = (C^{\sigma}_{\mathrm{new}})_i$, where

$$(C^{\sigma}_{\mathrm{new}})_i = \begin{cases} \mathbf{E}^{\sigma}_{i+1,1}\begin{bmatrix} \sqrt[n]{(\mathbf{R}^{\sigma}_{2,1})^T} & -\frac{1}{n}\left(\sqrt[n]{(\mathbf{R}^{\sigma}_{2,1})^T}\right)^{n-i+2}\mathbf{p}^{\sigma}_{2,1} \\ \mathbb{0}_3^T & 1 \end{bmatrix}(\mathbf{E}^{\sigma}_{i+1,1})^{-1} & \text{if } i \neq 1, \\ \begin{bmatrix} \sqrt[n]{(\mathbf{R}^{\sigma}_{2,1})^T} & -\frac{1}{n}\sqrt[n]{(\mathbf{R}^{\sigma}_{2,1})^T}\mathbf{p}^{\sigma}_{2,1} \\ \mathbb{0}_3^T & 1 \end{bmatrix} & \text{if } i = 1. \end{cases}$$

We remark that the expression for $(C^{\sigma}_{\mathrm{new}})_i$ is analogous to Definition 1, with the quantities associated with $\mathbf{E}^{\sigma}_{2,1}$ substituted for those of $\mathbf{E}^{\sigma}_{1,n}$. If this relation holds, then an induction step using the sequence of cycles in (24) proves the theorem.

If $\mathrm{An}(\mathbf{E}^{\sigma}_{1,n}) = 0$, then $C^{\sigma}_i = I_4 = (C^{\sigma}_{\mathrm{new}})_i$ for all $i \in \{1,\ldots,n\}$, and the statement is trivial. Assume $\mathrm{An}(\mathbf{E}^{\sigma}_{1,n}) \neq 0$. Define $(C^{\sigma}_{\mathrm{new}})_i := ((R^{\sigma}_{\mathrm{new}})_i, (p^{\sigma}_{\mathrm{new}})_i)$, and assume the conditions of the theorem. Recall that the action of a rotation matrix on a vector which is aligned with its axis of rotation is trivial. Since $\widehat{\mathrm{Ax}(\mathbf{E}^{\sigma}_{i,i-1})}\,\mathrm{Tran}(\mathbf{E}^{\sigma}_{i,i-1}) = \mathbb{0}_3$ for all $i \in \{1,\ldots,n\}$ by assumption, we have

$$\sqrt[n]{(\mathbf{R}^{\sigma}_{i,i-1})^T}\,\mathrm{Tran}(\mathbf{E}^{\sigma}_{i,i-1}) = \mathrm{Tran}(\mathbf{E}^{\sigma}_{i,i-1}) \quad \text{for any } i \in \{1,\ldots,n\}.$$

Further, results from [30, Chapters 2–3] state the following: if $E = (R,p) \in SE(3)$ such that $\widehat{\mathrm{Ax}(E)}\,\mathrm{Tran}(E) = \mathbb{0}_3$, then $p = \mathrm{An}(E)\,\mathrm{Tran}(E)$, and thus

$$E = \exp\left(\mathrm{An}(E)\widehat{\begin{bmatrix} \mathrm{Ax}(E) \\ \mathrm{Tran}(E) \end{bmatrix}}\right) = \begin{bmatrix} R & \mathrm{An}(E)\,\mathrm{Tran}(E) \\ \mathbb{0}_3^T & 1 \end{bmatrix}.$$

It follows that

$$\begin{aligned}
C_i^\sigma &= \mathbf{E}_{i+1,n}^\sigma \begin{bmatrix} \sqrt[n]{(\mathbf{R}_{1,n}^\sigma)^T} & -\frac{1}{n}\left(\sqrt[n]{(\mathbf{R}_{1,n}^\sigma)^T}\right)^{n-i+1} \mathbf{p}_{1,n}^\sigma \\ \mathbb{0}_3^T & 1 \end{bmatrix} (\mathbf{E}_{i+1,n}^\sigma)^{-1} \\
&= \mathbf{E}_{i+1,n}^\sigma \begin{bmatrix} \sqrt[n]{(\mathbf{R}_{1,n}^\sigma)^T} & -\frac{1}{n}\left(\sqrt[n]{(\mathbf{R}_{1,n}^\sigma)^T}\right)^{n-i+1} \operatorname{An}(\mathbf{E}_{1,n}^\sigma)\operatorname{Tran}(\mathbf{E}_{1,n}^\sigma) \\ \mathbb{0}_3^T & 1 \end{bmatrix} (\mathbf{E}_{i+1,n}^\sigma)^{-1} \\
&= \mathbf{E}_{i+1,n}^\sigma \begin{bmatrix} \sqrt[n]{(\mathbf{R}_{1,n}^\sigma)^T} & -\frac{1}{n}\operatorname{An}(\mathbf{E}_{1,n}^\sigma)\operatorname{Tran}(\mathbf{E}_{1,n}^\sigma) \\ \mathbb{0}_3^T & 1 \end{bmatrix} (\mathbf{E}_{i+1,n}^\sigma)^{-1} \\
&= \mathbf{E}_{i+1,n}^\sigma \exp\left(-\frac{1}{n}\operatorname{An}(\mathbf{E}_{1,n}^\sigma)\widehat{\begin{bmatrix} \operatorname{Ax}(\mathbf{E}_{1,n}^\sigma) \\ \operatorname{Tran}(\mathbf{E}_{1,n}^\sigma) \end{bmatrix}}\right)(\mathbf{E}_{i+1,n}^\sigma)^{-1}.
\end{aligned}$$

Applying Lemma 9, we have

$$\begin{aligned}
&\mathbf{E}_{i+1,n}^\sigma \exp\left(-\frac{1}{n}\operatorname{An}(\mathbf{E}_{1,n}^\sigma)\widehat{\begin{bmatrix} \operatorname{Ax}(\mathbf{E}_{1,n}^\sigma) \\ \operatorname{Tran}(\mathbf{E}_{1,n}^\sigma) \end{bmatrix}}\right)(\mathbf{E}_{i+1,n}^\sigma)^{-1} \\
&= \mathbf{E}_{i+1,n}^\sigma \exp\left(-\frac{1}{n}\operatorname{An}(\mathbf{E}_{1,n}^\sigma)E_{\sigma_1}\widehat{\begin{bmatrix} \operatorname{Ax}(\mathbf{E}_{2,1}^\sigma) \\ \operatorname{Tran}(\mathbf{E}_{2,1}^\sigma) \end{bmatrix}}E_{\sigma_1}^{-1}\right)(\mathbf{E}_{i+1,n}^\sigma)^{-1} \\
&= \mathbf{E}_{i+1,1}^\sigma \exp\left(-\frac{1}{n}\operatorname{An}(\mathbf{E}_{2,1}^\sigma)\widehat{\begin{bmatrix} \operatorname{Ax}(\mathbf{E}_{2,1}^\sigma) \\ \operatorname{Tran}(\mathbf{E}_{2,1}^\sigma) \end{bmatrix}}\right)(\mathbf{E}_{i+1,1}^\sigma)^{-1} \\
&= \mathbf{E}_{i+1,1}^\sigma \begin{bmatrix} \sqrt[n]{(\mathbf{R}_{2,1}^\sigma)^T} & -\frac{1}{n}\operatorname{An}(\mathbf{E}_{2,1}^\sigma)\operatorname{Tran}(\mathbf{E}_{2,1}^\sigma) \\ \mathbb{0}_3^T & 1 \end{bmatrix} (\mathbf{E}_{i+1,1}^\sigma)^{-1} \\
&= \mathbf{E}_{i+1,1}^\sigma \begin{bmatrix} \sqrt[n]{(\mathbf{R}_{2,1}^\sigma)^T} & -\frac{1}{n}\left(\sqrt[n]{(\mathbf{R}_{2,1}^\sigma)^T}\right)^{n-i+2} \mathbf{p}_{2,1}^\sigma \\ \mathbb{0}_3^T & 1 \end{bmatrix} (\mathbf{E}_{i+1,1}^\sigma)^{-1} \\
&= (C_{\text{new}}^\sigma)_i.
\end{aligned}$$

To see the last equality for the case when $i = 1$, we simply note that

$$\begin{aligned}
&= \mathbf{E}_{2,1}^\sigma \begin{bmatrix} \sqrt[n]{(\mathbf{R}_{2,1}^\sigma)^T} & -\frac{1}{n}\left(\sqrt[n]{(\mathbf{R}_{2,1}^\sigma)^T}\right)^{n-i+2} \mathbf{p}_{2,1}^\sigma \\ \mathbb{0}_3^T & 1 \end{bmatrix} (\mathbf{E}_{2,1}^\sigma)^{-1} \\
&= \begin{bmatrix} \sqrt[n]{(\mathbf{R}_{2,1}^\sigma)^T} & -\frac{1}{n}\left(\sqrt[n]{(\mathbf{R}_{2,1}^\sigma)^T}\right)^{n-i+2} \mathbf{p}_{2,1}^\sigma \\ \mathbb{0}_3^T & 1 \end{bmatrix},
\end{aligned}$$

since $\operatorname{Ax}(\mathbf{E}_{2,1}^\sigma)$ is aligned with $\mathbf{p}_{2,1}^\sigma = \operatorname{An}(\mathbf{E}_{2,1}^\sigma)\operatorname{Tran}(\mathbf{E}_{2,1}^\sigma)$ by assumption. □

We will need the following result for the proof of Theorem 4.

LEMMA 10 (rotational optimality). *Let $R \in SO(3)$. A global minimizer of the optimization problem*

$$\begin{aligned}
\textit{minimize} \quad & \operatorname{An}(R_1)^2 + \operatorname{An}(R_2)^2 \\
\textit{subject to} \quad & R_1 R_2 = R, \\
& R_1, R_2 \in SO(3),
\end{aligned}$$

is given by $R_1 = R_2 = \tilde{R}$, where $\tilde{R}$ is such that $\tilde{R}\tilde{R} = R$ and $\operatorname{An}(\tilde{R}) = \frac{1}{2}\operatorname{An}(R)$.

Similarly, if $p \in \mathbb{R}^3$*, then a global minimizer of the optimization problem*

$$\begin{aligned} \textit{minimize} \quad & \|p_1\|^2 + \|p_2\|^2 \\ \textit{subject to} \quad & p_1 + p_2 = p, \\ & p_1, p_2 \in \mathbb{R}^3, \end{aligned}$$

is given by $p_1 = p_2 = p/2$.

Proof. Let $R_1, R_2 \in SO(3)$ such that $R_1 R_2 = R$. It is trivial to show that if both $\mathrm{An}(R_1) > \mathrm{An}(R)$ and $\mathrm{An}(R_2) > \mathrm{An}(R)$, then R_1, R_2 are not optimal (compare this to the case where $R_1 = R_2 = \tilde{R}$). Therefore, assume without loss of generality that $\mathrm{An}(R_1) \leqslant \mathrm{An}(R)$. A consequence of the triangle inequality for the distance (6) implies that $\mathrm{An}(R_1) + \mathrm{An}(R_2) \geqslant \mathrm{An}(R) = \mathrm{An}(R_1) + (\mathrm{An}(R) - \mathrm{An}(R_1))$, and thus $\mathrm{An}(R_2) \geqslant (\mathrm{An}(R) - \mathrm{An}(R_1))$. We then have

$$\begin{aligned} \mathrm{An}(R_1)^2 + \mathrm{An}(R_2)^2 &\geqslant \mathrm{An}(R_1)^2 + (\mathrm{An}(R) - \mathrm{An}(R_1))^2 \\ &= \left(\frac{1}{2}\,\mathrm{An}(R) - \delta\right)^2 + \left(\frac{1}{2}\,\mathrm{An}(R) + \delta\right)^2 \\ &\geqslant \frac{1}{2}\left(\mathrm{An}(R)\right)^2, \end{aligned}$$

where $\delta = \frac{1}{2}\,\mathrm{An}(R) - \mathrm{An}(R_1)$. Notice that $2\,\mathrm{An}(\tilde{R})^2 = \frac{1}{2}\,\mathrm{An}(R)^2$. Therefore, $R_1 = R_2 = \tilde{R}$ is a global minimizer of the specified optimization.

The proof of the translational case is analogous. □

Proof of Theorem 4. Both optimizations (9) and (10) are feasible, as C^σ is a feasible solution (Theorem 2).

Case I *(rotation).* Assume $E_i = (R_i, \mathbb{0}_3)$ for all $i \in \{1, \dots, n\}$, and suppose that $C_1, C_2, \dots, C_n$ is a global minimizer of (9) (such a global minimizer exists by compactness of $SO(3)^n$). We prove the statement by showing that C_i must be equal to C_i^σ for all $i \in \{1, \dots, n\}$. We can apply a similarity transformation to each C_i for $i \in \{1, \dots, n\}$ to obtain a new set of matrices $(C_{\sigma_1})_{\mathrm{shift}}, (C_{\sigma_2})_{\mathrm{shift}}, \dots, (C_{\sigma_n})_{\mathrm{shift}}$ that satisfy $E_{\sigma_1} C_{\sigma_1} E_{\sigma_2} C_{\sigma_2} \cdots E_{\sigma_n} C_{\sigma_n} = \mathbf{E}^\sigma_{1,n} (C_{\sigma_1})_{\mathrm{shift}} (C_{\sigma_2})_{\mathrm{shift}} \cdots (C_{\sigma_n})_{\mathrm{shift}}$. The value of the cost function of (9) with respect to $(C_1)_{\mathrm{shift}}, (C_2)_{\mathrm{shift}}, \dots, (C_n)_{\mathrm{shift}}$ is the same as the cost function with respect to $C_1, C_2, \dots, C_n$; i.e., $\sum_{i=1}^n d(C_i, I_4)^2 = \sum_{i=1}^n d((C_i)_{\mathrm{shift}}, I_4)^2$. Indeed, if $C := (R_C, \mathbb{0}_3), E := (R, \mathbb{0}_3) \in SE(3)$ are arbitrary matrices of pure rotation, then $E C_i E^{-1} = (R R_C R^T, \mathbb{0}_3)$ and thus $d(C_i, I_4) = d(E C_i E^{-1}, I_4)$ since rotation angles are invariant under conjugation.

By assumption of feasibility, $(C_{\sigma_1})_{\mathrm{shift}} (C_{\sigma_1})_{\mathrm{shift}} \cdots (C_{\sigma_n})_{\mathrm{shift}} = \left(\mathbf{E}^\sigma_{1,n}\right)^{-1}$, which is a matrix of pure rotation. We deduce that for the solution $C_1, C_2, \dots, C_n$ to be optimal, $\mathrm{An}\left((C_1)_{\mathrm{shift}}\right) = \mathrm{An}\left((C_2)_{\mathrm{shift}}\right) = \cdots = \mathrm{An}\left((C_n)_{\mathrm{shift}}\right)$. Indeed, if this were not so, there would exist indices $i, i+1$ such that $\mathrm{An}\left((C_{\sigma_i})_{\mathrm{shift}}\right) \neq \mathrm{An}((C_{\sigma_{i+1}})_{\mathrm{shift}})$, and according to Lemma 10 we could lower the cost function by finding a matrix C that satisfies the equation $CC = (C_{\sigma_i})_{\mathrm{shift}} (C_{\sigma_{i+1}})_{\mathrm{shift}}$ and redefining our corrections accordingly. Further, a consequence of the triangle inequality tells us that in order to satisfy $(C_{\sigma_1})_{\mathrm{shift}} (C_{\sigma_1})_{\mathrm{shift}} \cdots (C_{\sigma_n})_{\mathrm{shift}} = \left(\mathbf{E}^\sigma_{1,n}\right)^{-1}$, where all of our corrections are pure rotations, we must have $\sum_{i=1}^n \mathrm{An}((C_i)_{\mathrm{shift}}) \geqslant \mathrm{An}(\mathbf{E}^\sigma_{1,n})$. These properties allow us to deduce $(C_{\sigma_i})_{\mathrm{shift}} = (C^\sigma_{\mathrm{EA}})_i$ for all $i \in \{1, \dots, n\}$, as this is the only choice such that $\mathrm{An}((C_1)_{\mathrm{shift}}) = \mathrm{An}((C_2)_{\mathrm{shift}}) = \cdots = \mathrm{An}((C_n)_{\mathrm{shift}}) = (1/n)\,\mathrm{An}(\mathbf{E}^\sigma_{1,n})$ (the invariance property of Theorem 3 holds in this setting). It follows that $C_i = C_i^\sigma$ for all $i \in \{1, \dots, n\}$ as desired.

Case II *(translation).* Now assume $E_i = (I_3, p_i)$ for all $i \in \{1, \ldots, n\}$. A global minimizer of (10) exists by the following logic: if $\mathrm{Cost}(C^\sigma)$ denotes the cost of C^σ, then the minimizer of (10) over the compact set consisting of pure translational correction matrices with translational components having norm less than or equal to $\mathrm{Cost}(C^\sigma)$ is a global minimizer, since any solution outside of this set has a cost larger than $\mathrm{Cost}(C^\sigma)$. Suppose that $C_1, C_2, \ldots, C_n$ is such a minimizer. We prove the statement by showing that C_i must be equal to C_i^σ for all $i \in \{1, \ldots, n\}$. Notice that any C_i remains unchanged under similarity transformation by any composition of E_i's:

$$C_i = \begin{bmatrix} I_3 & p \\ \mathbb{0}_3^T & 1 \end{bmatrix} \exp\left(\mathrm{An}(C_i)\widehat{\begin{bmatrix} \mathbb{0}_3 \\ \mathrm{Tran}(C_i) \end{bmatrix}}\right)\begin{bmatrix} I_3 & -p \\ \mathbb{0}_3^T & 1 \end{bmatrix}.$$

Therefore, we have $E_{\sigma_1}C_{\sigma_1}E_{\sigma_2}C_{\sigma_2}\cdots E_{\sigma_n}C_{\sigma_n} = \mathbf{E}_{1,n}^\sigma C_{\sigma_1}C_{\sigma_2}\cdots C_{\sigma_n}$. By assumption of optimality, we deduce that $\mathrm{An}(C_1) = \mathrm{An}(C_2) = \cdots = \mathrm{An}(C_n)$. Indeed, if this were not the case, then there would exist indices $i, i+1$ such that $\mathrm{An}(C_{\sigma_i}) \neq \mathrm{An}(C_{\sigma_{i+1}})$, and thus by Lemma 10 we could lower the cost function by replacing p_{σ_i} and $p_{\sigma_{i+1}}$ with $(p_{\sigma_i} + p_{\sigma_{i+1}})/2$. Further, by the triangle inequality, we must have $\sum_{i=1}^n \mathrm{An}(C_i) \geqslant \mathrm{An}(\mathbf{E}_{1,n}^\sigma)$. It follows that choosing $C_{\sigma_i} = (C_{\mathrm{EA}}^\sigma)_i$ for all $i \in \{1, \ldots, n\}$ is optimal, since this is the only choice such that $\|p_1\| = \|p_2\| = \cdots = \|p_n\| = \frac{1}{n}\|\mathbf{p}_{1,n}^\sigma\|$ (the invariance property of Theorem 3 holds). It follows that $C_i = C_i^\sigma$ for all $i \in \{1, \ldots, n\}$. ∎

Appendix B. Proof of Theorem 8. In this appendix, we assume the problem setup and notation of sections 5 and 6. We also make use of the bi-invariant Riemannian metric on $SO(3)$ obtained by projection of the metric (5); i.e., at any $R \in SO(3)$, we adopt[5] the inner product $\langle \cdot, \cdot \rangle_R : T_R SO(3) \times T_R SO(3) \to \mathbb{R}$ defined by $\langle R\widehat{v}, R\widehat{w} \rangle_R = v^T w$ and the associated geodesic distance function $d : SO(3) \times SO(3) \to [0, \pi]$ defined by $d(R_1, R_2) = \mathrm{An}(R_1^T R_2)$. This metric extends to $SO(3)^n$ in a natural way: for each $R := (R_1, R_2, \ldots, R_n) \in SO(3)^n$, define the inner product $\langle \cdot, \cdot \rangle_R : T_R SO(3)^n \times T_R SO(3)^n \to \mathbb{R}$ with the rule

$$\langle (R_1\widehat{v}_1, R_2\widehat{v}_2, \ldots, R_n\widehat{v}_n), (R_1\widehat{w}_1, R_2\widehat{w}_2, \ldots, R_n\widehat{w}_n) \rangle_R = \sum_{i=1}^n \langle R_i\widehat{v}_i, R_i\widehat{w}_i \rangle_{R_i}.$$

In accordance with notation of section 6, given a semicycle σ, a vector $R := (R_1, R_2, \ldots, R_n) \in SO(3)^n$, and a vector $p := (p_1, p_2, \ldots, p_n) \in \mathbb{R}^{3n}$, let $R_\sigma := \overline{R_{\sigma_1}}^\sigma \overline{R_{\sigma_2}}^\sigma \cdots \overline{R_{\sigma_{|\sigma|}}}^\sigma$ and p_σ be defined as the translational component of $\overline{E_{\sigma_1}}^\sigma \overline{E_{\sigma_2}}^\sigma \cdots \overline{E_{\sigma_{|\sigma|}}}^\sigma$, where $E_i := (R_i, p_i)$ for each $i \in \{1, \ldots, n\}$. For convenience, we also define the set $U := \{R \in SO(3)^n \mid \mathrm{An}(R_\sigma) = \pi \text{ for some } \sigma \in \mathcal{L}_f(G)\}$. Since $|\mathcal{L}_f(G)|$ is finite, it is trivial to show that the set U is closed in $SO(3)^n$: if $R \notin U$, then there exists a neighborhood that is not contained in U. Therefore, the notion of a derivative is well defined at points in the open set $SO(3)^n \backslash U$.

For a semicycle σ, define the map $e_\sigma^{\mathrm{rot}} : SO(3)^n \backslash U \to [0, \pi)$ by the rule $e_\sigma^{\mathrm{rot}}(R) = \mathrm{An}(R_\sigma)$. Further, given a vector $R \in SO(3)^n$, we define the map $e_{\sigma,R}^{\mathrm{tran}} : \mathbb{R}^{3n} \to \mathbb{R}_{\geqslant 0}$ by the rule $e_{\sigma,R}^{\mathrm{tran}}(p) = \|p_\sigma\|$. We write $(e_\sigma^{\mathrm{rot}})^2$, $(e_{\sigma,R}^{\mathrm{tran}})^2$ to denote the maps $R \mapsto (e_\sigma^{\mathrm{rot}}(R))^2$ and $p \mapsto (e_{\sigma,R}^{\mathrm{tran}}(p))^2$, respectively. We proceed with a few useful results.

LEMMA 11 (differentiability). *The maps* $(e_\sigma^{\mathrm{rot}})^2$ *and* $(e_{\sigma,R}^{\mathrm{tran}})^2$ *are differentiable.*

Proof. The mapping $e_{\sigma,R}^{\mathrm{tran}}$ is linear, and thus it is clear that $(e_{\sigma,R}^{\mathrm{tran}})^2$ is differentiable. Let $R := (R_1, R_2, \ldots, R_n) \in SO(3)^n \backslash U$. Since $(e_\sigma^{\mathrm{rot}})^2(R)$ is the squared rotation

[5]Without loss of generality we take $\alpha = 1$.

angle of a composition of rotations, differentiability is clear at any point such that $R_\sigma \neq I_3$. Suppose $R_\sigma = I_3$, and let $\omega := (\omega_1, \omega_2, \dots, \omega_n) \in \mathbb{R}^{3n}$. Let $\delta > 0$ be small and define a curve $\gamma : (-\delta, \delta) \to SO(3)^n$ by

$$\gamma(t) = (R_1 \exp(t\widehat{\omega}_1), R_2 \exp(t\widehat{\omega}_2), \dots, R_n \exp(t\widehat{\omega}_n)).$$

Notice that $\frac{d\gamma}{dt}\ |_{t=0} = (R_1\widehat{\omega}_1, R_2\widehat{\omega}_2, \dots, R_n\widehat{\omega}_n)$, and thus it suffices to show that $\frac{d}{dt}(e_\sigma^{\text{rot}})^2(\gamma(\cdot))\ |_{t=0}$ is well defined for any choice of the vector ω. We have

$$(e_\sigma^{\text{rot}})^2(\gamma(t)) = \text{An}\left(\overline{R_{\sigma_1} \exp(t\widehat{\omega}_{\sigma_1})}^\sigma\, \overline{R_{\sigma_2} \exp(t\widehat{\omega}_{\sigma_2})}^\sigma \cdots \overline{R_{\sigma_{|\sigma|}} \exp(t\widehat{\omega}_{\sigma_{|\sigma|}})}^\sigma\right)^2.$$

By appropriate conjugations of the exponential terms, one obtains

$$(e_\sigma^{\text{rot}})^2(\gamma(t)) = \text{An}\left(R_\sigma \exp(t\widehat{\nu}_{\sigma_1}) \exp(t\widehat{\nu}_{\sigma_2}) \cdots \exp(t\widehat{\nu}_{\sigma_{|\sigma|}})\right)^2,$$

where $\nu_{\sigma_i} \in \mathbb{R}^3$, $\|\nu_{\sigma_i}\| = \|\omega_{\sigma_i}\|$ for all $i \in \{1, \dots, |\sigma|\}$ (by invariance of rotation angles under conjugation). For $t > 0$ sufficiently small, by the triangle inequality we have

$$\begin{aligned}(e_\sigma^{\text{rot}})^2(\gamma(t)) &= \text{An}\left(R_\sigma \exp(t\widehat{\nu}_{\sigma_1}) \exp(t\widehat{\nu}_{\sigma_2}) \cdots \exp(t\widehat{\nu}_{\sigma_{|\sigma|}})\right)^2 \\ &\leqslant (\text{An}(R_\sigma) + t|\sigma|)^2 = O(t^2) \quad \text{as } t \to 0,\end{aligned}$$

since $\text{An}(R_\sigma) = 0$ by assumption. It follows that $\frac{d}{dt}(e_\sigma^{\text{rot}})^2(\gamma(\cdot))\ |_{t=0} = \lim_{t\to 0} \frac{1}{t} O(t^2) = 0$. The statement of Lemma 11 follows. ∎

LEMMA 12 (gradient behavior). *Let* $R := (R_1, R_2, \dots, R_n) \in SO(3)^n \backslash U$ *and* $p := (p_1, p_2, \dots, p_n) \in \mathbb{R}^{3n}$. *The gradient*

$$\nabla(e_\sigma^{\text{rot}})^2(R) := ((\nabla(e_\sigma^{\text{rot}})^2(R))_1, (\nabla(e_\sigma^{\text{rot}})^2(R))_2, \dots, (\nabla(e_\sigma^{\text{rot}})^2(R))_n) \in T_R SO(3)^n$$

is given by

$$(\nabla(e_\sigma^{\text{rot}})^2(R))_i = \begin{cases} -2R_i \,\text{An}(R_\sigma)\widehat{\text{Ax}(C^\sigma_{j(\sigma,i)})} & \text{if } \mathcal{E}_i \in \sigma, \\ \widehat{\mathbb{0}}_3 & \text{otherwise,} \end{cases}$$

where $j(\sigma, i)$ *is the index such that* $\sigma_{j(\sigma,i)} = i$. *If, in addition,* $R_\sigma = I_3$, *then* $\nabla(e_{\sigma,R}^{\text{tran}})^2(p) := ((\nabla(e_{\sigma,R}^{\text{tran}})^2(p))_1, \dots, (\nabla(e_{\sigma,R}^{\text{tran}})^2(p))_n) \in \mathbb{R}^{3n}$ *is given by*

$$(\nabla(e_{\sigma,R}^{\text{tran}})^2(p))_i = \begin{cases} -2R_i \|p_\sigma\| \,\text{Tran}(C^\sigma_{j(\sigma,i)}) & \text{if } \mathcal{E}_i \in \sigma, \\ \mathbb{0}_3 & \text{otherwise.} \end{cases}$$

Proof. The vectors $\nabla(e_\sigma^{\text{rot}})^2(R)$ and $\nabla(e_{\sigma,R}^{\text{tran}})^2(p)$ exist by Lemma 11. Let $i \in \{1, \dots, n\}$. If $\mathcal{E}_i \notin \sigma$, it is trivial to show that $(\nabla(e_{\sigma,R}^{\text{tran}})^2(p))_i = \mathbb{0}_3$ and $(\nabla(e_\sigma^{\text{rot}})^2(R))_i = \widehat{\mathbb{0}}_3$. Thus, in what follows assume that $\mathcal{E}_i \in \sigma$. We break the proof into two parts.

Part I *(gradient of* $(e_\sigma^{\text{rot}})^2$*).* If $\text{An}(R_\sigma) = 0$, then R is a local minimizer of e_σ^{rot}, and thus the statement about $\nabla(e_\sigma^{\text{rot}})^2$ is clear. Assume $\text{An}(R_\sigma) \neq 0$, let $\delta > 0$ be small, and define $\gamma : (-\delta, \delta) \to SO(3)^{|\sigma|}$ by

$$\gamma(t) = \left(R_{\sigma_1} \exp\left(-t|\sigma|^{-\frac{1}{2}} \widehat{\text{Ax}(C_1^\sigma)}\right), \dots, R_{\sigma_{|\sigma|}} \exp\left(-t|\sigma|^{-\frac{1}{2}} \widehat{\text{Ax}(C_{|\sigma|}^\sigma)}\right)\right).$$

It is easily verified that $\frac{d\gamma}{dt}|_{t=0} = (-|\sigma|^{-\frac{1}{2}} R_{\sigma_1}\widehat{\text{Ax}(C_1^\sigma)}, \dots, -|\sigma|^{-\frac{1}{2}} R_{\sigma_{|\sigma|}}\widehat{\text{Ax}(C_{|\sigma|}^\sigma)})$ and has unit magnitude with respect to the natural metric associated with the inner product defined on $T_R SO(3)^{|\sigma|}$. We now calculate the derivative in the direction of $\frac{d\gamma}{dt}|_{t=0}$. For $t > 0$ small we have

$$\begin{aligned}&(e_\sigma^{\text{rot}})^2(\gamma(t))\\ &= \text{An}\left(\overline{R_{\sigma_1}\exp\left(-t|\sigma|^{-\frac{1}{2}}\widehat{\text{Ax}(C_1^\sigma)}\right)}^\sigma \cdots \overline{R_{\sigma_{|\sigma|}}\exp\left(-t|\sigma|^{-\frac{1}{2}}\widehat{\text{Ax}(-tC_{|\sigma|}^\sigma)}\right)}^\sigma\right)^2\\ &= \text{An}\left(R_\sigma \exp\left(-t|\sigma|^{-\frac{1}{2}}\widehat{\text{Ax}\left((C_{\text{EA}}^\sigma)_1\right)}\right)\exp\left(-t|\sigma|^{-\frac{1}{2}}\widehat{\text{Ax}\left((C_{\text{EA}}^\sigma)_{|\sigma|}\right)}\right)\right)^2.\end{aligned}$$

By construction, $\text{Ax}\left((C_{\text{EA}}^\sigma)_i\right) = -\text{Ax}(R_\sigma)$ for every i such that $\mathcal{E}_i \in \sigma$, and thus for $t > 0$ sufficiently small, $(e_\sigma^{\text{rot}})^2(\gamma(t)) = (\text{An}(R_\sigma) + t\sqrt{|\sigma|})^2$ and thus $\frac{d}{dt}(e_\sigma^{\text{rot}})^2(\gamma(\cdot))|_{t=0} = \lim_{t\to 0}\frac{1}{t}(\text{An}(R_\sigma) + t\sqrt{|\sigma|})^2 - \text{An}(R_\sigma)^2 = 2\sqrt{|\sigma|}\,\text{An}(R_\sigma)$. From the manipulations above, it is also clear that for any other curve $\tilde{\gamma} : (-\delta,\delta) \to SO(3)^{|\sigma|}$ with a derivative of unit magnitude at $t = 0$, if $\frac{d\tilde{\gamma}}{dt}\,|_{t=0} \neq \frac{d\gamma}{dt}\,|_{t=0}$, then $\frac{d}{dt}(e_\sigma^{\text{rot}})^2(\tilde{\gamma}(\cdot))\,|_{t=0} < \frac{d}{dt}(e_\sigma^{\text{rot}})^2(\gamma(\cdot))\,|_{t=0}$. Indeed, using similar manipulations and letting $t > 0$ be small, we find $(e_\sigma^{\text{rot}})^2(\tilde{\gamma}(t)) = \text{An}(R_\sigma\tilde{R})^2$ for some $\tilde{R} \in SO(3)$ with an axis of rotation that does not point in the same direction as $\text{Ax}(R_\sigma)$, and thus $\text{An}(R_\sigma\tilde{R})^2 < (\text{An}(R_\sigma) + t\sqrt{\sigma})^2$. By definition of the gradient we have $(\nabla(e_\sigma^{\text{rot}})^2(R))_i = -2R_i\,\text{An}(R_\sigma)\widehat{\text{Ax}(C_{j(\sigma,i)}^\sigma)}$.

Part II *(gradient of* $(e_{\sigma,R}^{\text{tran}})^2$*)*. For this part of the proof, we assume that $R_\sigma = I_3$. If $\|p_\sigma\| = 0$, then $\text{Tran}(C_i^\sigma) = \mathbb{0}_3$, and since p is a local minimizer of the function $e_{\sigma,R}^{\text{tran}}$, the statement about $\nabla(e_{\sigma,R}^{\text{tran}})^2$ holds. Assume $\|p_\sigma\| \neq 0$. Then, C_i^σ represents pure translation for each i, and thus by (4) the vector $\text{Tran}(C_i^\sigma)$ is a unit vector that points in the direction of the translational component of C_i^σ. Similarly to the rotation case, for some small $\delta > 0$ we define a curve $\gamma : (-\delta,\delta) \to \mathbb{R}^{3|\sigma|}$ by $\gamma(t) = (p_{\sigma_1} - t|\sigma|^{-\frac{1}{2}}R_{\sigma_1}\text{Tran}(C_1^\sigma), \ldots, p_{\sigma_{|\sigma|}} - t|\sigma|^{-\frac{1}{2}}R_{\sigma_{|\sigma|}}\text{Tran}(C_{|\sigma|}^\sigma))$. We now find the derivative of $(e_{\sigma,R}^{\text{tran}})^2$ in the direction $\frac{d\gamma}{dt}\,|_{t=0}$. Let $t > 0$ be small. Note that for any choice of i such that $\mathcal{E}_i \in \sigma$, we have

(25)
$$\overline{\begin{bmatrix} R_{\sigma_i} & p_{\sigma_i} - t|\sigma|^{-\frac{1}{2}}R_{\sigma_i}\text{Tran}(C_i^\sigma) \\ \mathbb{0}_3^T & 1\end{bmatrix}}^\sigma = \overline{\begin{bmatrix} R_{\sigma_i} & p_{\sigma_i} \\ \mathbb{0}_3^T & 1\end{bmatrix}\begin{bmatrix} I_3 & -t|\sigma|^{-\frac{1}{2}}\text{Tran}(C_i^\sigma) \\ \mathbb{0}_3^T & 1\end{bmatrix}}^\sigma,$$

and thus finding the value of $(e_{\sigma,R}^{\text{tran}})^2(\gamma(t))$ is equivalent to finding the squared norm of the translational component of the composition

$$\overline{E_{\sigma_1}\tilde{C}_1^\sigma}^\sigma\,\overline{E_{\sigma_2}\tilde{C}_2^\sigma}^\sigma \cdots \overline{E_{\sigma_{|\sigma|}}\tilde{C}_{|\sigma|}^\sigma}^\sigma = \overline{E_{\sigma_1}E_{\sigma_2}}^\sigma \cdots \overline{E_{\sigma_{|\sigma|}}}^\sigma(\tilde{C}_{\text{EA}}^\sigma)_1(\tilde{C}_{\text{EA}}^\sigma)_2\cdots(\tilde{C}_{\text{EA}}^\sigma)_{|\sigma|},$$

where $\tilde{C}_i^\sigma := (I_3, -t|\sigma|^{-\frac{1}{2}}\text{Tran}(C_i^\sigma)), (\tilde{C}_{\text{EA}}^\sigma)_i := (I_3, -t|\sigma|^{-\frac{1}{2}}\text{Tran}((C_{\text{EA}}^\sigma)_i)) \in SE(3)$. By construction, $\text{Tran}((C_{\text{EA}}^\sigma)_i) = -p_\sigma/\|p_\sigma\|$; thus

$$(e_{\sigma,R}^{\text{tran}})^2(\gamma(t)) = \left\|p_\sigma + t\sqrt{|\sigma|}\frac{p_\sigma}{\|p_\sigma\|}\right\|^2 = (\|p_\sigma\| + t|\sigma|^{\frac{1}{2}})^2,$$

and

$$\frac{d}{dt}(e_{\sigma,R}^{\text{tran}})^2(\gamma(\cdot))\bigg|_{t=0} = \lim_{t\to 0}\frac{1}{t}(\|p_\sigma\| + t\sqrt{|\sigma|})^2 - \|p_\sigma\|^2 = 2\sqrt{|\sigma|}\|p_\sigma\|.$$

By arguments similar to those in the rotation case, we can argue that for any other curve $\tilde{\gamma} : (-\delta,\delta) \to \mathbb{R}^{|\sigma|n}$ with a derivative of unit magnitude at $t = 0$, if $\frac{d\tilde{\gamma}}{dt}\,|_{t=0}\neq$

$\frac{d\gamma}{dt}\mid_{t=0}$, then $\frac{d}{dt}(e^{\mathrm{tran}}_{\sigma,R})^2(\tilde{\gamma}(\cdot))\mid_{t=0}\leqslant \frac{d}{dt}(e^{\mathrm{tran}}_{\sigma,R})^2(\gamma(\cdot))\mid_{t=0}$. Thus we have deduced that

$$(\nabla(e^{\mathrm{tran}}_{\sigma,R})^2(p))_i = -2R_i\|p_\sigma\|\,\mathrm{Tran}(C^\sigma_{j(\sigma,i)}). \qquad \square$$

We are now poised to complete the proof of Theorem 8.

Proof of Theorem 8. We consider two cases.

Case I *(there exists $\sigma \in \mathcal{L}_f(G)$ such that $R_\sigma \neq I_3$).* In this setting, any solution to the equal angle flow with given initial measurements will not reach a configuration in which $R_\sigma = I_3$ for all $\sigma \in \mathcal{L}_f(G)$ in finite time, since the derivatives of all R_i asymptotically approach 0 as the system approaches any such configuration. As such, the translational component of any solution will be constant, and thus we need only study the evolution of the rotational components for a complete understanding of the solution curves. Consider the candidate Lyapunov function $V : SO(3)^n\backslash U \to \mathbb{R}_{\geqslant 0}$ given by

$$V(R) = \sum_{\sigma\in\mathcal{L}_f(G)} \frac{1}{|\sigma|}\left(\ln(\pi^2) - \ln\left(\pi^2 - (e^{\mathrm{rot}}_\sigma)^2(R)\right)\right).$$

Note that this function is differentiable as it is the composition of differentiable functions (Lemma 11) and that $V(R) \to \infty$ as $R \to U$. Differentiating termwise and using Lemma 12,

$$\nabla V(R) = \sum_{\sigma\in\mathcal{L}_f(G)} \frac{1}{|\sigma|(\pi^2 - (e^{\mathrm{rot}}_\sigma)^2(R))}\nabla(e^{\mathrm{rot}}_\sigma)^2(R),$$

and thus we have

$$(\nabla V(R))_i = -2R_i \sum_{\substack{\sigma\in\mathcal{L}_f(G)\\ \mathcal{E}_i\in\sigma}} \frac{\mathrm{An}(R_\sigma)}{|\sigma|(\pi^2 - \mathrm{An}(R_\sigma)^2)}\widehat{\mathrm{Ax}(C^\sigma_{j(\sigma,i)})},$$

where $i \in \{1,\dots,n\}$. It follows that under the equal angle flow, $\dot{R}_1,\dots,\dot{R}_n$ is always in the direction of the negative gradient of the function V. Thus, if a solution exists, it would have the property that $\langle\nabla V(R), \dot{R}\rangle_R \leqslant 0$ for all time $t \geqslant 0$. Since $V \to \infty$ as $R \to U$, it follows that any solution lies entirely in a compact set which is disjoint from U. The vector field in Definition 7 is easily shown to be locally Lipschitz continuous, as it is the composition of locally Lipschitz continuous mappings, and thus existence and uniqueness of solutions to the equal angle flow follow immediately (see, e.g., [1, Chapter 4]). Invoking Lasalle's invariance principle (see, e.g., [12, Chapter 6]), we conclude that solution curves will converge to a set such that $\langle\nabla V(R), \dot{R}\rangle_R = 0$ (or, equivalently, the set where $\nabla V(R)$ vanishes).

To complete the proof, we note that, through our construction of the set $\mathcal{L}_f(G)$, each semicycle σ contains at least one edge that is not contained in any of the other semicycles. It follows that $\nabla V(R)$ does not vanish unless $\mathrm{An}(R_\sigma) = 0$ for all $\sigma \in \mathcal{L}_f(G)$. We conclude that under the equal angle flow, we obtain convergence to a set of rotations such that $R_\sigma = I_3$ for all $\sigma \in \mathcal{L}_f(G)$.

Case II *($R_\sigma \neq I_3$ for all $\sigma \in \mathcal{L}_f(G)$).* We now assume that the nominal measurements have the additional property that $\overline{E_{\sigma_1}(0)}^\sigma\,\overline{E_{\sigma_2}(0)}^\sigma\cdots\overline{E_{\sigma_{|\sigma|}}(0)}^\sigma$ represents pure translation for all $\sigma \in \mathcal{L}_f(G)$. In this setting, the rotational component of any solution to the equal angle flow will remain constant for all time, and thus R remains fixed. Assume $p(t) := (p_1(t), \ldots, p_n(t))$, where $p(0)$ is equal to our nominal translation measurements, and define the candidate Lyapunov function $V : \mathbb{R}^{3n} \to \mathbb{R}_{\geqslant 0}$ by

$$V(p) = \sum_{\sigma \in \mathcal{L}_f(G)} \frac{1}{|\sigma|} (e^{\mathrm{tran}}_{\sigma,R})^2(p).$$

Note that V is differentiable (Lemma 11). From Lemma 12, we have

$$(\nabla V(p))_i = \sum_{\substack{\sigma \in \mathcal{L}_f(G) \\ \mathcal{E}_i \in \sigma}} \frac{1}{|\sigma|} (\nabla (e^{\mathrm{tran}}_{\sigma,R})^2(p))_i = -2R_i \sum_{\substack{\sigma \in \mathcal{L}_f(G) \\ \mathcal{E}_i \in \sigma}} \frac{\|p_\sigma\|}{|\sigma|} \mathrm{Tran}(C^\sigma_{j(\sigma,i)}),$$

where $i \in \{1, \ldots, n\}$. Under the equal angle flow, $\dot{p}$ is in the direction of $-\nabla V(p)$. Thus, we have $(\nabla V(p))^T \dot{p}(t) \leqslant 0$ for all time $t \geqslant 0$. Using the same argument as in the rotational case, we conclude that there exists a well-defined, unique solution $\mathbb{R}_{\geqslant 0} \ni t \mapsto (E_1(t), E_2(t), \ldots, E_n(t))$. Since $\nabla V(p) \neq \mathbb{0}_{3n}$ unless $p_\sigma = \mathbb{0}_3$ for all $\sigma \in \mathcal{L}_f(G)$, the statement follows once again by Lasalle's invariance principle. ∎

REFERENCES

[1] R. Abraham, J. E. Marsden, and T. S. Ratiu, *Manifolds, Tensor Analysis, and Applications*, Appl. Math. Sci. 75, 2nd ed., Springer-Verlag, New York, 1988.

[2] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2007.

[3] P.-A. Absil and J. Malick, *Projection-like retractions on matrix manifolds*, SIAM J. Optim., 22 (2012), pp. 135–158.

[4] R. L. Adler, J. P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub, *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA J. Numer. Anal., 22 (2002), pp. 359–390.

[5] C. Altafini, *The De Casteljau algorithm on SE(3)*, in Nonlinear Control in the Year 2000, A. Isidori, F. Lamnabhi-Lagarrigue, and W. Respondek, eds., Springer, London, 2001, pp. 23–34.

[6] E. Amaldi, L. Liberti, and F. Maffioli, *Algorithms for finding minimum fundamental cycle bases in graphs*, Electron. Notes Discrete Math., 17 (2004), pp. 29–33.

[7] J. Aspnes, T. Eren, D. K. Goldenberg, A. S. Morse, W. Whiteley, Y. R. Yang, B. D. O. Anderson, and P. Belhumeur, *A theory of network localization*, IEEE Trans. Mobile Comput., 5 (2006), pp. 1663–1678.

[8] P. Barooah and J. P. Hespanha, *Estimation from relative measurements: Algorithms and scaling laws*, IEEE Control Syst. Mag., 27 (2007), pp. 57–74.

[9] P. Barooah and J. P. Hespanha, *Error scaling laws for linear optimal estimation from relative measurements*, IEEE Trans. Inform. Theory, 55 (2009), pp. 5661–5673.

[10] N. Boumal, A. Singer, P.-A. Absil, and V. D. Blondel, *Cramér–Rao bounds for synchronization of rotations*, Inf. Inference, 3 (2014), pp. 1–39.

[11] P. Brida, J. Duha, and M. Krasnovsky, *On the accuracy of weighted proximity based localization in wireless sensor networks*, in Personal Wireless Communications, IFIP—The International Federation for Information Processing 245, R. Bestak, B. Simak, and E. Kozlowska, eds., Springer, Boston, 2007, pp. 423–432.

[12] F. Bullo and A. D. Lewis, *Geometric Control of Mechanical Systems. Modeling, Analysis, and Design for Simple Mechanical Control Systems*, Springer-Verlag, New York, Heidelberg, Berlin, 2004.

[13] F. Bullo and R. M. Murray, *Proportional Derivative (PD) Control on the Euclidean Group*, Technical Report Caltech/CDS 95–010, California Institute of Technology, Pasadena, CA, 1995; available online at http://resolver.caltech.edu/CaltechCDSTR:1995.CIT-CDS-95-010.

[14] M. Cao, B. D. O. Anderson, and A. S. Morse, *Sensor network localization with imprecise distance measurements*, Systems Control Lett., 55 (2006), pp. 887–893.

[15] K. Daniilidis, *Hand-eye calibration using dual quaternions*, Internat. J. Robotics Res., 18 (1999), pp. 286–298.

[16] L. Doherty, K. S. J. Pister, and L. El Ghaoui, *Convex position estimation in wireless sensor networks*, in Proceedings of the Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies, Anchorage, AK, 2001, pp. 1655–1663.

[17] G. Dubbelman, P. Hansen, B. Browning, and M. B. Dias, *Orientation only loop-closing with closed-form trajectory bending*, in Proceedings of the IEEE International Conference on Robotics and Automation, Saint Paul, MN, 2012, pp. 815–821.

[18] S. Esquivel, F. Woelk, and R. Koch, *Calibration of a multi-camera rig from non-overlapping views*, in Pattern Recognition, F. A. Hamprecht, C. Schnörr, and B. Jähne, eds., Springer-Verlag, Berlin, Heidelberg, 2007, pp. 82–91.

[19] P. T. Fletcher, *Geodesic regression and the theory of least squares on Riemannian manifolds*, Int. J. Comput. Vis., 105 (2013), pp. 171–185.

[20] L. R. Foulds, *Graph Theory Applications*, Universitext, Springer-Verlag, New York, 1995.

[21] G. Galbiati and E. Amaldi, *On the approximability of the minimum fundamental cycle basis problem*, in Approximation and Online Algorithms, Lecture Notes in Comput. Sci. 2909, Springer, Berlin, 2004, pp. 151–164.

[22] C. Gramkow, *On averaging rotations*, J. Math. Imaging Vision, 15 (2001), pp. 7–16.

[23] V. Guillemin and A. Pollack, *Differential Topology*, AMS, Providence, RI, 2010.

[24] R. Hartley, J. Trumpf, Y. Dai, and H. Li, *Rotation averaging*, Int. J. Comput. Vis., 103 (2013), pp. 1–39.

[25] R. Horaud and F. Dornaika, *Hand-eye calibration*, Internat. J. Robotics Res., 14 (1995), pp. 195–210.

[26] D. Q. Huynh, *Metrics for 3D rotations: Comparison and analysis*, J. Math. Imaging Vision, 35 (2009), pp. 155–164.

[27] T. Kavitha, C. Liebchen, K. Mehlhorn, D. Michail, R. Rizzi, T. Ueckerdt, and K. A. Zweig, *Cycle bases in graphs characterization, algorithms, complexity, and applications*, Comput. Sci. Rev., 3 (2009), pp. 199–243.

[28] G. Mao, B. Fidan, and B. D. O. Anderson, *Wireless sensor network localization techniques*, Comput. Netw., 51 (2007), pp. 2529–2553.

[29] D. Moore, J. Leonard, D. Rus, and S. Teller, *Robust distributed network localization with noisy range measurements*, in Proceedings of the 2nd International Conference on Embedded Networked Sensor Systems, Baltimore, MD, ACM, New York, 2004, pp. 50–61.

[30] R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*, CRC Press, Boca Raton, FL, 1994.

[31] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, *Locating the nodes: Cooperative localization in wireless sensor networks*, IEEE Signal Processing Mag., 22 (2005), pp. 54–69.

[32] N. Patwari, A. O. Hero, and M. Perkins, *Relative location estimation in wireless sensor networks*, IEEE Trans. Signal Process., 51 (2003), pp. 2137–2148.

[33] G. Piovan, I. Shames, B. Fidan, F. Bullo, and B. D. O. Anderson, *On frame and orientation localization for relative sensing networks*, Automatica J. IFAC, 49 (2013), pp. 206–213.

[34] S. I. Roumeliotis and G. A. Bekey, *Distributed multirobot localization*, IEEE Trans. Robot. Autom., 18 (2002), pp. 781–795.

[35] G. C. Sharp, S. W. Lee, and D. K. Wehe, *Multiview registration of 3D scenes by minimizing error between coordinate frames*, IEEE Trans. Pattern Anal. Mach. Intell., 26 (2004), pp. 1037–1050.

[36] K. H. Strobl and G. Hirzinger, *Optimal hand-eye calibration*, in Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 2006, pp. 4647–4653.

[37] M. M. Sysło, *On cycle bases of a graph*, Networks, 9 (1979), pp. 123–132.

[38] R. Tron and R. Vidal, *Distributed 3-D localization of camera sensor networks from 2-D image measurements*, IEEE Trans. Automat. Control, 59 (2014), pp. 3325–3340.

[39] J. Wang, R. K. Ghosh, and S. K. Das, *A survey on sensor localization*, J. Control Theory Appl., 8 (2010), pp. 2–11.

[40] Y. Xu, Y. Ouyang, Z. Le, J. Ford, and F. Makedon, *Mobile anchor-free localization for wireless sensor networks*, in Distributed Computing in Sensor Systems, Lecture Notes in Comput. Sci. 4549, J. Aspnes, C. Scheideler, A. Arora, and S. Madden, eds., Springer-Verlag, Berlin, Heidelberg, 2007, pp. 96–109.

[41] M. Žefran, V. Kumar, and C. Croke, *On the generation of smooth three-dimensional rigid body motions*, IEEE Trans. Robot. Autom., 14 (1998), pp. 576–589.