The Dynamic Bearing Observability Matrix Nonlinear Observability and Estimation for Multi-Agent Systems

Fabrizio Schiano, Roberto Tron

The Dynamic Bearing Observability Matrix Nonlinear Observability and Estimation for Multi-Agent Systems

Fabrizio Schiano and Roberto Tron

***Abstract*— We consider the problem of localization in multi-agent formations with bearing only measurements, and analyze the fundamental observability properties for dynamic agents. The current well-established approach is based on the so-called *rigidity matrix*, and its algebraic properties (e.g., its rank and nullspace). This method is typically motivated using first-order derivatives, and shows, among other facts, that the *global scale* of the formation is not observable. This work shows that current results represent an incomplete view of the problem. In particular, we show that *1)* current methods are a particular instantiation of nonlinear observability theory, *2)* we can introduce the concept of the *dynamic bearing observability matrix* from higher order derivatives to study the observability of dynamic formations, and *3)* the global scale is, in fact, generally observable when the agents move according to known inputs. We use tools from Riemannian geometry and Lie group theory to tackle, in a general and principled way, the general formulation of the localization problem with states that include both rotations and translations. Finally, we verify our theoretical results by deriving and applying, in both simulations and real experiments on UAVs, a centralized Extended Kalman Filter on Lie groups that is able to estimate the global scale of a moving formation.**

I. Introduction

Multiagent systems, and in particular multirobot systems (e.g., quadrotors), have been the subject of research for more than thirty years. These systems have many potential advantages with respect to single-agent systems (e.g., decentralized processing, resilience against failures of individual agents, faster task completion times, etc.) A fundamental problem in this area is the one of *localization*, broadly defined as the process of determining the position of the agents in a common reference frame from a sparse set of relative measurements between them, and without the aid of an external centralized system (such as GPS). This problem becomes of paramount importance, for instance, when a team of robots has to navigate autonomously in an unknown environment, or needs to collaborate on a physical task (e.g., transporting a load [1]). In this case it is usually required for the robots to be able to localize themselves w.r.t. other agents of the formation. During recent years, a number of sensor modalities for such measurements have been considered. From a theoretical point of view, the most important aspect of these different modalities is the quantity of information they provide. For instance, the sensors could provide estimates of distance (e.g., from wireless signal strength), bearing directions (e.g., with monocular cameras), translations (e.g., with stereo or depth cameras), or both rotations and translations (e.g., using cameras and two-view Structure from Motion [2], [3]). In this paper, we consider the case of bearing measurements, that is, we assume that each agent can measure the direction (but not the distance) of a subset of neighboring agents in its own reference frame. This setup is arguably the most practical with today's most popular hardware, which, due to weight, cost, and power consumption considerations, is usually limited to an Inertial Measurement Unit (IMU) and a monocular camera [4]. Our goal is to show that, despite the very limited information provided by this type of measurements (direction of relative translations alone), it is possible to reconstruct the full 3-D pose (rotation and translation) of the agents, including the global scale, up to a global gauge ambiguity. We demonstrate this *1)* theoretically through an application of nonlinear observability analysis and Riemannian geometry, and *2)* practically by applying an Extended Kalman Filter in both simulations and experiments.

Prior work. The problem of localization from bearing-only measurements has appeared in a variety of domains, such as (to cite a few) sensor network localization [5]–[7] and formation control [8]–[10] in controls and robotics, Structure from Motion [11], [12] in computer vision, and graph drawings [13] in discrete mathematics. Most of the literature has focused on the development of distributed algorithms (especially in the sensor network and robotics communities), but centralized solutions have also been considered (mostly in the computer vision community). In this work, rather than specific algorithms, we are interested in analyzing the fundamental aspects of the localization problem. In this direction, for our case of interest involving bearing-only measurement, there has been a considerable amount of work for developing a *theory of rigidity* [14]–[19], which can predict what information can be recovered from the available measurements (i.e., whether the solution is "unique"). While most of these works considered only agents in 2-D, recent work has considered also the 3-D case [20]. The commonly accepted result is that, when the number and connectivity of the measurement graph is sufficiently high (that is, when the graph is *rigid*), then, *for static agents*, the solution to the bearing-only localization problem is unique up to a *rototranslation* and a *contraction/expansion* of the whole formation. This is determined by considering the nullspace of a so-called *rigidity matrix*. Most of the existing works, however, do not explicitly consider the case of *dynamic* agents. At a high level, one could expect that if the agents know their own velocities in their own local frames (e.g., because they control them, or measure them using the onboard IMU), then they could use this *metric* information to avoid the scale ambiguity. Recent works [21] have pursued this idea, but do not provide a full, rigorous analysis rooted in

F. Schiano is with Inria, Univ Rennes, CNRS, Campus de Beaulieu, 35042 Rennes Cedex, France fabrizio.schiano@inria.fr

R.Tron is with the Department of Mechanical Engineering at Boston University, Boston, MA 02215, http://sites.bu.edu/tron, tron@bu.edu

nonlinear observability analysis and Riemannian geometry. One disadvantage of [21] is the presence of two estimators in *series* for estimating the scale of the formation (one filter estimates the distances over a selection of edges, and the second filter recovers a correctly scaled estimation of the formation configuration). Stability of this cascaded structure is difficult to prove (indeed, nothing is said in [21]), while our algorithm achieves the same result with only one *single* EKF. On the other hand, [21] is able to determine the optimal motion for the agents in order to maximize the observability of the scale factor. We plan to exploit the ideas in [21] for a similar characterization in the context of our EKF estimation. Nonetheless, ideas related to this approach have been successfully explored in the context of single-agent Simultaneous Localization and Mapping (SLAM) [4], [22] and localization from distance measurements [23].

In this regard, nonlinear observability (the problem of determining if the state of a nonlinear dynamical system can be reconstructed by knowing its inputs and outputs) is a classical topic in automatic controls [24]–[26], and it now constitutes textbook material [27]. However, it has never been explicitly applied to the problem of localization from bearing-only measurements. On the other hand, Riemannian geometry has been applied in the context of geometric control and estimation of mechanical systems in general [28], [29], and quadrotors in particular [1], [30], [31]. A Riemannian geometry formulation has also been used for multiagent localization with unscaled relative poses [32], but it has never been applied together with nonlinear observability analysis to the bearing-only case.

Finally, in this paper we propose a validation of our theoretical results using an Extended Kalman Filter (EKF) for statistical filtering of states evolving on Riemannian manifolds. The main advantage of the EKF formulation is that it is relatively easy to derive; in fact, both centralized [33] and decentralized [34] implementations have been proposed specifically for multiagent systems, although without considering states evolving on Riemannian manifolds. On the other hand, it is known that the EKF is not optimal for nonlinear systems. Developing filtering techniques with optimality guarantees on Riemannian manifolds (and Lie groups in particular) is still an active field of research [35], [36]. Since the goal of the paper is to simply use filtering as a validation of the theoretical derivations, we opted for a straight (although suboptimal), centralized (as opposed to distributed) application of the EKF, albeit with the explicit consideration of the Riemannian geometry of the states.

Paper contributions. In this paper, we make several contributions to the state of the art:

- We show how the study of rigidity is, in fact, a particular instance of classical nonlinear observability analysis;
- Using this insight, we propose the notion of *dynamic bearing observability matrix* (DBOM), which extends the standard notion of rigidity matrix for the case moving agents with known inputs;
- By (numerically) analyzing the rank of the DBOM, we show that *the global scale of the formation is generally observable*;
- We show how tools from Riemannian geometry can be employed to carry out the observability analysis for states evolving on the space of rigid body motions;
- We derive and apply a centralized Extended Kalman Filter (EKF) on Riemannian manifolds that empirically verifies the theory (i.e., that shows that the global scale can be indeed recovered).

Overall, we show how bearing-only measurements (which, taken individually, do not contain any scale information) and local linear and angular velocity information can be used to recover the entire state of the agents up to a common rotation and translation (since all measurements are relative and do not have any relation to external reference systems, this last ambiguity appears to be unavoidable, even with dynamic agents). Moreover, by explicitly using the Riemannian geometry of the space of poses (which is based, among other elements, on the use of rotations matrices) throughout the paper (both for the observability analysis and the statistical filter), we avoid the problems given by other representations (e.g., the singularities of Euler angles, and the non-uniqueness of quaternions). Finally, in this paper we do not perform a full, analytical characterization of the nullspace of the DBOM, and we do not consider distributed filtering solutions. However, these are interesting future directions that are enabled by the present work.

Paper overview. The paper is organized as follows. Sect. II reviews notions from several areas that are necessary to carry out our analysis. In Sect. III we introduce the novel concept of *Dynamic Bearing Observability Matrix*. Sect. IV illustrates our EKF design. Finally, Sect. V reports experimental results with a group of quadrotor UAVs, followed by Sect. VI that concludes the paper and gives possible future directions.

II. Preliminaries

A. General notation

Let $\mathcal{W}$ represent an absolute 3-D world reference frame, and $\mathcal{A}_i$ represent a body reference frame attached to the i-th agent. We use $\{\boldsymbol{e}_m\}_{m=1}^3$ to denote the standard $\mathbb{R}^3$ basis. We also let $\mathbf{1}_N$ and $\boldsymbol{I}_N$ represent a vector of all ones and the identity matrix of dimension N, respectively. The operator $\mathrm{stack}(\cdot)$ returns a matrix containing a vertical stacking of the arguments.

B. Formation, Agent and Measurement Model

As customary, we model the formation of robots with a directed sensing graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the vertex set $\mathcal{V} = \{1 \dots N\}$ represents the agents, and the edge set $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ contains the pairs of agents $(i, j) \in \mathcal{E}$ for which agent j can be sensed from agent i. We denote as $\mathcal{N}_i = \{j \in \mathcal{V} | \, (i, \, j) \in \mathcal{E}\} \subset \mathcal{V}$ the set of neighbors of an agent i in $\mathcal{G}$.

We model the state of an agent $i \in \mathcal{V}$ as a pose $\boldsymbol{q}_i = (\boldsymbol{p}_i, \boldsymbol{R}_i)$, where $\boldsymbol{p}_i \in \mathbb{R}^3$ represents the translation of the origin of $\mathcal{A}_i$ expressed in $\mathcal{W}$, and $\boldsymbol{R}_i \in \mathrm{SO}(3)$ represents the rotation transforming directions from $\mathcal{A}_i$ to $\mathcal{W}$. We denote the space of rigid poses as $\mathrm{SE}(3)$ (the detailed definition of $\mathrm{SO}(3)$ and its geometry is postponed to Sect. II-C). We assume a simple first order model for the 6-D dynamics of each agent

$$\dot{\boldsymbol{q}}_i = \left(\dot{\boldsymbol{p}}_i, \dot{\boldsymbol{R}}_i\right) = \left(\boldsymbol{R}_i \boldsymbol{v}_i, \boldsymbol{R}_i \hat{\boldsymbol{w}}_i\right) \tag{1}$$

$$= \sum_{k=1}^{3} \left(\boldsymbol{R}_i \boldsymbol{e}_k, 0\right) v_{i_k} + \sum_{k=1}^{3} \left(0, \boldsymbol{R}_i \hat{\boldsymbol{e}}_k\right) w_{i_k}, \qquad (2)$$

where $\boldsymbol{v}_i, \boldsymbol{w}_i \in \mathbb{R}^3$ represent, respectively, the linear and angular velocities expressed in $\mathcal{A}_i$, and v_{i_k}, w_{i_k} represent their components along the $\boldsymbol{e}_k$ basis vector. We use this model for generality, but the results of this work could be easily specialized to other cases (e.g., considering only the positions of the agents, or only the 2-D yaw angle, as done in the majority of previous works).

We assume that each robot is equipped with a sensor (onboard calibrated camera) that allows it to measure the *relative bearing vector* w.r.t. an agent $j \in \mathcal{N}_i$ in its own reference frame $\mathcal{A}_i$, i.e., the 3-D unit-norm vector

$$\boldsymbol{\beta}_{ij} = \boldsymbol{h}_{ij}(\boldsymbol{q}_i, \boldsymbol{q}_j) = \boldsymbol{R}_i^T \frac{\boldsymbol{p}_j - \boldsymbol{p}_i}{\|\boldsymbol{p}_j - \boldsymbol{p}_i\|} = \boldsymbol{R}_i^T \frac{\boldsymbol{p}_{ij}}{d_{ij}} \in \mathbb{S}^2, \quad (3)$$

where $\boldsymbol{p}_{ij} = \boldsymbol{p}_j - \boldsymbol{p}_i$, $d_{ij} = \|\boldsymbol{p}_j - \boldsymbol{p}_i\|$.

As in [16], [17], [21], [37], we assume that we have only available the inputs $\{\boldsymbol{v}_i, \boldsymbol{w}_i\}_{i\in\mathcal{V}}$, and the measurements $\{\boldsymbol{\beta}_{ij}\}_{(i,j)\in\mathcal{E}}$. In particular, we have access neither to the absolute states $\boldsymbol{q}_i$, nor to the global reference frame $\mathcal{W}$. Throughout the nonlinear observability analysis we will refer to the different components of the vector $\boldsymbol{\beta}_{ij}$ as:

$$\beta_{ij_m} = \boldsymbol{e}_m^T \boldsymbol{\beta}_{ij} \in \mathbb{R},\ m \in \{1, 2, 3\}. \qquad (4)$$

Remark 1. *While we will individually consider each one of the three elements of each bearing $\boldsymbol{\beta}_{ij}$ as a separate output, in reality the fact that $\boldsymbol{\beta}_{ij} \in \mathbb{S}^2$ implies that only two outputs are algebraically independent. The effect of this is that the bearing rigidity matrix that we will derive will contain more rows than strictly needed (i.e., some rows will be automatically linearly dependent). However, this does not change the result of the rank-based observability test.*

C. Elements of Riemannian geometry

This section covers the basic Riemannian geometry notions that are used in the derivations below. We will be mostly concerned with three manifolds: the Euclidean space $\mathbb{R}^3$, the space of 3-D rotations $\mathrm{SO}(3) = \{\boldsymbol{R} \in \mathbb{R}^{3\times3} : \boldsymbol{R}^{\mathrm{T}}\boldsymbol{R} = I, \det(\boldsymbol{R}) = 1\}$, the space of 3-D poses $\mathrm{SE}(3) = \{(\boldsymbol{p}, \boldsymbol{R}) : \boldsymbol{R} \in \mathrm{SO}(3), \boldsymbol{p} \in \mathbb{R}^3\}$, and the space of N 3-D poses $\mathrm{SE}(3)^N$. These manifold are in fact *Lie groups*, but we will not make use of this fact.

a) Tangent spaces: We denote as $T_x\mathcal{M}$ the *tangent space* of a manifold $\mathcal{M}$ at a point $x \in \mathcal{M}$. The tangent space at a point can be identified as the vector space spanned by the tangents of the curves passing through that point; for instance, if $\boldsymbol{R}(t) : I \to \mathrm{SO}(3)$ is a parametrized curve in $\mathrm{SO}(3)$ defined on some interval $I \subset \mathbb{R}$ around zero, then $\dot{\boldsymbol{R}}(0) \in T_{\boldsymbol{R}(0)}\mathrm{SO}(3)$. For $\mathbb{R}^3$, the tangent space at each point can be identified with $\mathbb{R}^3$ itself. For $\mathrm{SO}(3)$ however, we first need to redefine the usual *hat* $\hat{(\cdot)}$ and *vee* $(\cdot)^\vee$ operators between $\mathbb{R}^3$ and the set of skew-symmetric matrices in $\mathbb{R}^{3\times3}$ as follows, with $\boldsymbol{v} = [v_1, v_2, v_3]^T$:

$$\boldsymbol{v}^{\boldsymbol{R}\wedge} = \boldsymbol{R}\begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} = \boldsymbol{R}\hat{\boldsymbol{v}},\ \boldsymbol{V}^{\boldsymbol{R}\vee} = \left(\boldsymbol{R}^T\boldsymbol{V}\right)^\vee \quad (5)$$

where $\boldsymbol{V}$ is any vector $\boldsymbol{V} \in T_{\boldsymbol{R}}\mathrm{SO}(3)$[1]. It can be shown that the tangent space of $\mathrm{SO}(3)$ is given by

$$T_{\boldsymbol{R}}\mathrm{SO}(3) = \{\boldsymbol{R}\hat{\boldsymbol{v}} : \boldsymbol{v} \in \mathbb{R}^3\}. \qquad (6)$$

We define a basis for $T_{\boldsymbol{R}}\mathrm{SO}(3)$ as $\{\boldsymbol{e}_m^{\boldsymbol{R}\vee}\}_{m=1}^3$; it follows that any vector $\boldsymbol{V} \in T_{\boldsymbol{R}}\mathrm{SO}(3)$ can be expressed as a vector $\boldsymbol{v} \in \mathbb{R}^3$ of *local coordinates* in this basis with the relation $\boldsymbol{v} = \boldsymbol{V}^{\boldsymbol{R}\vee}$. For instance, for the curve $\boldsymbol{R}(t)$ defined above, letting $\boldsymbol{R}_0 = \boldsymbol{R}(0)$ we will have $\dot{\boldsymbol{R}}(0) = \boldsymbol{w}^{\boldsymbol{R}_0\wedge}$ for some $\boldsymbol{w} \in \mathbb{R}^3$, or, equivalently, $\boldsymbol{w} = \dot{\boldsymbol{R}}(0)^{\boldsymbol{R}_0\vee}$. In the case where $\boldsymbol{R}(t)$ represents a physical time-varying rotation, using the convention given in Sec. II-B, the vector $\boldsymbol{w}$ coincides with a vector of angular velocities expressed in the body frame (see also (1)). The tangent space of $T_{\boldsymbol{q}}\mathrm{SE}(3)$ can be identified with the direct sum $\mathbb{R}^3 \oplus T_{\boldsymbol{R}}\mathrm{SO}(3)$ (i.e., a tangent for $\mathrm{SE}(3)$ is simply a tangent for $\mathbb{R}^3$ together with a tangent for $\mathrm{SO}(3)$). Similarly, the tangent space of $\mathrm{SE}(3)^N$ is simply the direct sum of N copies of $T_{\boldsymbol{q}}\mathrm{SE}(3)$. A representation in local coordinates of a vector $T_{\boldsymbol{q}}\mathrm{SE}(3)$ can be obtained by stacking the local coordinate representation of each rotational component (as discussed above), with the translational components.

Remark 2. *Note that $\dot{\boldsymbol{q}}$ can be interpreted as a tuple of linear and angular velocities or as a single vector expressed directly in local coordinates.*

b) Riemannian metrics: A Riemannian metric $\langle , \rangle$ smoothly assigns an inner product to each tangent space. The standard Riemannian metric for $\mathbb{R}^3$ is the usual inner product. The standard Riemannian metric for $\mathrm{SO}(3)$ is defined as

$$\langle \boldsymbol{R}\hat{\boldsymbol{v}}_1, \boldsymbol{R}\hat{\boldsymbol{v}}_2 \rangle = \frac{1}{2}\operatorname{tr}(\hat{\boldsymbol{v}}_1^{\mathrm{T}}\hat{\boldsymbol{v}}_2) = \boldsymbol{v}_1^{\mathrm{T}}\boldsymbol{v}_2, \qquad (7)$$

where $\boldsymbol{R}\hat{\boldsymbol{v}}_1, \boldsymbol{R}\hat{\boldsymbol{v}}_2 \in T_{\boldsymbol{R}}\mathrm{SO}(3)$ are two tangent vectors. For $\mathrm{SE}(3)$, we use the metric given by the sum of the two previous metrics, and for $\mathrm{SE}(3)^N$, the sum of the metrics for each copy of $\mathrm{SE}(3)$.

c) Gradients and how to compute them: The *gradient* of a differentiable function $f(x)$, $f : \mathcal{M} \to \mathbb{R}$ computed at a point x_0 on a manifold $\mathcal{M}$ is defined as the unique tangent vector $\nabla_x f(x_0)$ such that, for all curves $x(t)$ with $x(0) = x_0$,

$$\langle \nabla_x f(x_0), \dot{x}(0) \rangle = \frac{\mathrm{d}}{\mathrm{d}t} f\big(x(t)\big)\Big|_{t=0}. \qquad (8)$$

For $\mathbb{R}^3$, it can be shown that this definition coincides with the more common definition as a vector of partial derivatives. For $\mathrm{SO}(3)$, we can use (8) to compute gradients of any arbitrary function f in a few steps. First, we consider a fictitious parametrized curve $\boldsymbol{R}(t)$ in $\mathrm{SO}(3)$. Then we use the chain rule to compute $\frac{\mathrm{d}}{\mathrm{d}t} f\big(\boldsymbol{R}(t)\big)$, the derivative of the function along the curve. It can be shown that this derivative (when it exists) can always be written as

$$\frac{\mathrm{d}}{\mathrm{d}t} f(\boldsymbol{R}) = \operatorname{tr}(\boldsymbol{M}^{\mathrm{T}}\dot{\boldsymbol{R}}) = \operatorname{tr}\big(\operatorname{skew}(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{M})^{\mathrm{T}}\boldsymbol{R}^{\mathrm{T}}\dot{\boldsymbol{R}}\big), \quad (9)$$

where $\boldsymbol{M}$ is some matrix in $\mathbb{R}^{3\times3}$ which in general depends on $\boldsymbol{R}$, $\operatorname{skew}(\boldsymbol{A}) = \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^{\mathrm{T}})$ is an operator that extracts

[1] If $\boldsymbol{R}$ is not present in the superscript, or if the $\boldsymbol{R} = \boldsymbol{I}$, the definitions of *hat* and *vee* operators are the classical ones present in the literature.

the skew-symmetric component of a matrix, and the explicit dependency on t has been omitted for brevity. The first equality in (9) comes from the fact that the differential of a map (gradients are a particular case of differentials) are always linear maps [38], and they can be expressed as linear functionals using the trace operator [39]. The second equality in (9) comes from the characterization of $T_{\boldsymbol{R}}\mathrm{SO}(3)$ given in (6) and the fact that $\mathrm{tr}(\boldsymbol{S}\hat{\boldsymbol{v}}) = 0$ for any symmetric matrix $\boldsymbol{S} \in \mathbb{R}^{3\times3}$ and $\boldsymbol{v} \in \mathbb{R}^3$. Comparing (9) with (8), we then obtain $\nabla_{\boldsymbol{R}} f(\boldsymbol{R}_0) = 2\boldsymbol{R}\,\mathrm{skew}(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{M})$ (this is equivalent to the formula given, e.g., in [40]); in local coordinates, this becomes $2\,\mathrm{skew}(\boldsymbol{R}^{\mathrm{T}}\boldsymbol{M})^{\vee}$. Informally, we refer to this set of steps as the *trace trick*. For $\mathrm{SE}(3)$ and $\mathrm{SE}(3)^N$, the computation of the gradient reduces to a separate computation for each component.

Remark 3. *In this paper we represent rotations using rotation matrices. Compared to other representations (such as Euler angles or quaternions), this representation is unambiguous, does not have singularities, and, as shown above, provides a relatively straightforward way to compute gradients (see, e.g., [31] for additional insights).*

D. Elements of local nonlinear observability

Let $\boldsymbol{q} = (\boldsymbol{q}_1, \dots, \boldsymbol{q}_N) \in \mathrm{SE}(3)^N$ (according to remark 2), $\boldsymbol{\beta} = \mathrm{stack}(\{\boldsymbol{\beta}_{ij}\}) \in \mathbb{R}^{3|\mathcal{E}|}$, and the *input vector fields* $\boldsymbol{g}_{v_{i_k}}, \boldsymbol{g}_{w_{i_k}}$ on $\mathrm{SE}(3)^N$ which are obtained by appropriately padding with zeros the corresponding vectors in (2). Then, the dynamical model of the entire network can be considered as a nonlinear system with affine inputs:

$$\begin{aligned} \dot{\boldsymbol{q}} &= \sum_{i\in\mathcal{V}}\sum_{k=1}^{3}(\boldsymbol{g}_{v_{i_k}} v_{i_k} + \boldsymbol{g}_{w_{i_k}} w_{i_k}), \\ \boldsymbol{\beta} &= \boldsymbol{h}(\boldsymbol{q}), \end{aligned} \tag{10}$$

where $\boldsymbol{h} : \mathrm{SE}(3)^N \to \mathbb{R}^{3|\mathcal{E}|}$, $\boldsymbol{h} = \mathrm{stack}(\{\boldsymbol{h}_{ij}\})$.

The goal of nonlinear observability theory [22] applied to problem (10) is to determine what parts of the state $\boldsymbol{q}$ can be reconstructed from the outputs $\boldsymbol{\beta}$ and the inputs $\{\boldsymbol{v}_i, \boldsymbol{w}_i\}$. Following the notation of [22], we indicate the kth order *Lie derivative* of a function h along the vector fields $\boldsymbol{f}_1, \dots, \boldsymbol{f}_k$ as $L^k_{\boldsymbol{f}_1,\dots,\boldsymbol{f}_k} h$. The definition of Lie derivative is given by letting $L^0 h = h$, and then, recursively:

$$\mathrm{d}L^k_{\boldsymbol{f}_1,\dots,\boldsymbol{f}_k} h = \nabla_{\boldsymbol{q}} L^k_{\boldsymbol{f}_1,\dots,\boldsymbol{f}_k} h, \tag{11}$$

$$L^{k+1}_{\boldsymbol{f}_1,\boldsymbol{f}_2,\dots,\boldsymbol{f}_{k+1}} h = \langle \mathrm{d}L^k_{\boldsymbol{f}_1,\dots,\boldsymbol{f}_k} h, \boldsymbol{f}_{k+1}\rangle, \tag{12}$$

where $\langle , \rangle$ is the Riemannian metric on $\mathrm{SE}(3)^N$ described in Sect. II-C, and $\mathrm{d}L^k_{\boldsymbol{f}_1,\dots,\boldsymbol{f}_k} h$ is a shorthand notation for the gradient of a Lie derivative. In nonlinear observability, the function h is set to be an output of the system, and the vector fields $\boldsymbol{f}_1, \dots, \boldsymbol{f}_k$ are taken to be the input vector fields.

In our case, we consider each β_{ijm} (that is, each element of $\boldsymbol{\beta}$) as a separate output of the system. In order to carry out the local nonlinear observability analysis at a particular configuration $\boldsymbol{q}$, it is necessary to define the subspace

$$\mathrm{d}\Omega = \mathrm{span}\big(\{\mathrm{d}L^k_{\boldsymbol{f}_1,\dots,\boldsymbol{f}_k}\beta_{ijm}\}\big), \tag{13a}$$

$$\text{where } k \in \{0, 1, \dots\}, (i,j) \in \mathcal{E}, m \in \{1,2,3\}, \tag{13b}$$

$$\boldsymbol{f}_1, \dots, \boldsymbol{f}_k \in \{\boldsymbol{g}_{v_{i_k}}, \boldsymbol{g}_{w_{i_k}}\}. \tag{13c}$$

As discussed in [27, Sect. 1.9], it is sufficient to consider Lie derivatives up to the order $k = 6N - 1$ (where $6N$ is the dimension of the system (10)). In practice, we will numerically verify that $k = 1$ is already sufficient under general conditions to show that the global scale of the system can be recovered. We define $\mathrm{d}\Omega^{\perp}$ to be the *annihilator* of $\mathrm{d}\Omega$, that is, the subspace of $T_{\boldsymbol{q}}\mathrm{SE}(3)^N$ such that

$$\langle v, n\rangle = 0 \quad \forall v \in \mathrm{d}\Omega, n \in N \tag{14}$$

The annihilator tells us the *locally unobservable modes* of the system [22], [27], that is, what variations of the state $\boldsymbol{q}$ cannot be observed under any choice of the inputs (and their Lie brackets). In practice, to find $\mathrm{d}\Omega^{\perp}$ and compute its dimension, we need to switch to a local coordinate representation. To avoid introducing additional notation, we redefine $\mathrm{d}L^k_{\boldsymbol{f}_1,\dots,\boldsymbol{f}_k} h$ to be the vector in $\mathbb{R}^{6N}$ of local coordinates (as opposed to an abstract tangent vector in $T_{\boldsymbol{q}}\mathrm{SE}(3)^N$). Similarly, we redefine $\mathrm{d}\Omega$ as a matrix (the original subspace $\mathrm{d}\Omega$ is given by the row span of this matrix):

$$\mathrm{d}\Omega = \mathrm{stack}\big(\{\mathrm{d}L^k_{\boldsymbol{f}_1,\dots,\boldsymbol{f}_k}\beta_{ijm}\}\big), \tag{15}$$

where the indexes are the same as in (13). Intuitively, the matrix $\mathrm{d}\Omega$ generalizes the classical observability matrix used for linear systems [24], [25], [27]. Finally, the annihilator $\mathrm{d}\Omega^{\perp}$ is redefined to be the nullspace of $\mathrm{d}\Omega$, $\mathrm{d}\Omega^{\perp} = \mathrm{null}(\mathrm{d}\Omega)$.

In the next section we will give the details of the computation of $\mathrm{d}\Omega$ for Lie derivatives of order up to $k = 1$ for our system (10).

III. *Dynamic* Bearing Observability Matrix

In this section we introduce the notion of *Dynamic Bearing Observability Matrix*[2] (DBOM). We define the DBOM $\tilde{\boldsymbol{R}}$ to be equal to the matrix $\mathrm{d}\Omega$ computed with the gradients of Lie derivatives of order up to $k = 1$. More explicitly:

$$\tilde{\boldsymbol{R}} = \mathrm{stack}(\{\mathrm{d}L^0\beta_{ijm}\}, \{\mathrm{d}L^1_{\boldsymbol{f}_1}\beta_{ijm}\}) = \mathrm{stack}(\tilde{\boldsymbol{R}}_A, \tilde{\boldsymbol{R}}_B), \tag{16}$$

where the matrix $\tilde{\boldsymbol{R}}_A$ and $\tilde{\boldsymbol{R}}_B$ contain the gradients of the Lie derivative of order, respectively, $k = 0$ and $k = 1$, and will be described in detail in Sects. III-A and III-B.

We anticipate here that the matrix $\tilde{\boldsymbol{R}}_A$ is equivalent to the traditional *bearing rigidity matrix* first introduced in [41], and then expanded upon in various papers (e.g., [17], [18], [20]). However, here we *1)* derive it for full 6-D states in $\mathrm{SE}(3)$, using a clear interpretation with respect to the Riemannian geometry of the space, and *2)* give it the interpretation of a first step in a full nonlinear observability analysis. Intuitively, since this matrix includes only zeroth order Lie derivatives, its properties tell us which parts of the state can be estimated without providing any input (static agents). As expect from previous works, and compatibly with the intuition, global scaling (contraction/expansion) of the formation generate tangent vectors that are in the nullspace of $\tilde{\boldsymbol{R}}_A$, meaning that the global scale is not observable.

The matrix $\tilde{\boldsymbol{R}}_B$ constitutes the main novelty in our analysis. Intuitively, since this matrix includes first order Lie derivatives,

[2]We chose this name to stress that the scale of a formation based on bearing measurements is retrievable through *dynamic* information.

its properties tell us what can be estimated by moving the agents with constant inputs. As we will numerically verify in the following, this matrix contributes to reducing the dimension of the nullspace of $\tilde{\boldsymbol{R}}$ by one. The direction that is removed corresponds exactly to the contraction/expansion motion. This is compatible with the intuition above: when agents move, there is a parallax effect that can be exploited to get an estimate of the unknown scales. In the remainder of this section we include only the main results of our analysis. Please refer to [42] for more detailed derivations.

A. Matrix $\tilde{\boldsymbol{R}}_A$

In order to compute the matrix $\tilde{\boldsymbol{R}}_A$, since the zeroth-order Lie derivatives $L^0\beta_{ijm}$ are simply equal to the function themselves, we can directly focus on computing their gradients $\nabla_{\boldsymbol{q}} L^0\beta_{ijm}$. For this purpose, we will use the trace-trick method. Assuming that $\boldsymbol{\beta}_{ij}$ moves along a fictitious curve $\boldsymbol{\beta}_{ij}(t)$, we compute the following:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}L^0\beta_{ijm} = \langle\nabla_{\boldsymbol{p}_i}L^0\beta_{ijm},\dot{\boldsymbol{p}}_i\rangle + \langle\nabla_{\boldsymbol{p}_j}L^0\beta_{ijm},\dot{\boldsymbol{p}}_j\rangle + \\ +\langle\nabla_{\boldsymbol{R}_i}L^0\beta_{ijm},\dot{\boldsymbol{R}}_i\rangle + \langle\nabla_{\boldsymbol{R}_j}L^0\beta_{ijm},\dot{\boldsymbol{R}}_j\rangle = \boldsymbol{e}_m^T\dot{\boldsymbol{\beta}}_{ij}\end{aligned} \tag{17}$$

Expanding $\boldsymbol{e}_m^T\dot{\boldsymbol{\beta}}_{ij}$, and after some tedious algebra, relying on the steps detailed in Sect. II-C, we can extract the desired gradients (in local coordinates):

$$\nabla_{\boldsymbol{p}_i}L^0\left(\boldsymbol{e}_m^T\boldsymbol{\beta}_{ij}\right) = -\frac{1}{d_{ij}}\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\boldsymbol{R}_i^T, \tag{18}$$

$$\nabla_{\boldsymbol{p}_j}L^0\left(\boldsymbol{e}_m^T\boldsymbol{\beta}_{ij}\right) = +\frac{1}{d_{ij}}\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\boldsymbol{R}_i^T, \tag{19}$$

$$\nabla_{\boldsymbol{R}_i}L^0\left(\boldsymbol{e}_m^T\boldsymbol{\beta}_{ij}\right) = -\left(\mathrm{skew}\left(2\boldsymbol{\beta}_{ij}\boldsymbol{e}_m^T\right)^\vee\right)^T, \tag{20}$$

$$\nabla_{\boldsymbol{R}_j}L^0\left(\boldsymbol{e}_m^T\boldsymbol{\beta}_{ij}\right) = \boldsymbol{0}, \tag{21}$$

where $\boldsymbol{P}_{ij} = \boldsymbol{I}_3 - \boldsymbol{\beta}_{ij}\boldsymbol{\beta}_{ij}^T$ is the orthogonal projector onto the orthogonal complement of $\boldsymbol{\beta}_{ij}$. Equations (18)–(21) then give the $(1\times 6N)$ k-th row block of $\tilde{\boldsymbol{R}}_A$ associated to the edge $(i,j)\in\mathcal{E}$, of the form

$$\left[\ \boldsymbol{-0-}\ \ -\frac{\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\boldsymbol{R}_i^T}{d_{ij}}\ \ \left(-\mathrm{skew}\left(2\boldsymbol{\beta}_{ij}\boldsymbol{e}_m^T\right)^\vee\right)^T\ \ \boldsymbol{-0-}\ \ \frac{\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\boldsymbol{R}_i^T}{d_{ij}}\ \ \boldsymbol{-0-}\ \ \boldsymbol{-0-}\ \right], \tag{22}$$

where the blocks are ordered following the ordering of the translational and rotational states in $\boldsymbol{q}$. Notice that, in this formula, this block row is expressed in the frame $\mathcal{W}$. It is also possible to express it in the local frame using the same ideas as [18].

B. Matrix $\tilde{\boldsymbol{R}}_B$

In order to compute the matrix $\tilde{\boldsymbol{R}}_B$, it is necessary to first compute the first-order Lie derivatives $L^1_{\boldsymbol{f}_1}\beta_{ijm} = \langle\nabla_{\boldsymbol{q}}L^0\beta_{ijm},\boldsymbol{f}_1\rangle$, where $\boldsymbol{f}_1\in\{\boldsymbol{g}_{v_{i_k}},\boldsymbol{g}_{w_{i_k}}\}_{\imath\in\mathcal{V}}$. Note that, for a given i,j, all these Lie derivative are zero except for $\boldsymbol{f}_1\in\{\boldsymbol{g}_{v_{i_k}},\boldsymbol{g}_{w_{i_k}},\boldsymbol{g}_{v_{j_k}}\}$ (for $\boldsymbol{f}_1=\boldsymbol{g}_{w_{j_k}}$, the Lie derivative is zero due to (21)). We can therefore consider only the latter ones. For instance, let us focus on the case $\boldsymbol{f}_1=\boldsymbol{g}_{v_{i_k}}$. To compute the gradient of $L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm}$ we can employ again the trace-trick method.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm} = \langle\nabla_{\boldsymbol{p}_i}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm},\dot{\boldsymbol{p}}_i\rangle + \langle\nabla_{\boldsymbol{p}_j}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm},\dot{\boldsymbol{p}}_j\rangle + \\ +\langle\nabla_{\boldsymbol{R}_i}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm},\dot{\boldsymbol{R}}_i\rangle + \langle\nabla_{\boldsymbol{R}_j}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm},\dot{\boldsymbol{R}}_j\rangle\end{aligned} \tag{23}$$

Notice that, however, while expanding above we will obtain terms that depend on $\dot{\boldsymbol{\beta}}_{ij}$. Similarly to what we mentioned in Sect. II-C, one can show that this dependency is linear. More explicitly, we can rewrite (23) as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm} = \langle\bar{\nabla}_{\boldsymbol{p}_i}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm},\dot{\boldsymbol{p}}_i\rangle + \langle\bar{\nabla}_{\boldsymbol{p}_j}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm},\dot{\boldsymbol{p}}_j\rangle + \\ +\langle\bar{\nabla}_{\boldsymbol{R}_i}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm},\dot{\boldsymbol{R}}_i\rangle + \langle\bar{\nabla}_{\boldsymbol{R}_j}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm},\dot{\boldsymbol{R}}_j\rangle + \boldsymbol{K}_{ij}\dot{\boldsymbol{\beta}}_{ij},\end{aligned} \tag{24}$$

where $\boldsymbol{K}_{ij}$ is a matrix in $\mathbb{R}^{3\times 3}$. We can exploit (24) to compute the rank of the DBOM $\tilde{\boldsymbol{R}}$, while avoiding the explicit computation of the terms in $\boldsymbol{K}_{ij}$, thus simplifying the analytical expressions involved; this is because we can collect all the $\{\boldsymbol{K}_{ij}\}_{(i,j)\in\mathcal{E}}$ into a $3|\mathcal{E}|\times 3|\mathcal{E}|$ matrix $\boldsymbol{K} = \mathrm{diag}\big(\{\boldsymbol{K}_{ij}\}\big)$, and then rewrite (16) as

$$\tilde{\boldsymbol{R}} = \begin{bmatrix}\tilde{\boldsymbol{R}}_A \\ \tilde{\boldsymbol{R}}_C + \boldsymbol{K}\tilde{\boldsymbol{R}}_A\end{bmatrix}, \tag{25}$$

where $\tilde{\boldsymbol{R}}_C$ is defined in the same way as $\tilde{\boldsymbol{R}}_B$, but by using the modified (and analytically simpler) gradients $\bar{\nabla}_{\boldsymbol{q}}$ instead of the full gradients $\nabla_{\boldsymbol{q}}$. Eq. (25) implies that

$$\mathrm{rank}\begin{bmatrix}\tilde{\boldsymbol{R}}_A \\ \tilde{\boldsymbol{R}}_B\end{bmatrix} = \mathrm{rank}\left(\begin{bmatrix}\boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{K} & \boldsymbol{I}\end{bmatrix}\begin{bmatrix}\tilde{\boldsymbol{R}}_A \\ \tilde{\boldsymbol{R}}_C\end{bmatrix}\right) = \mathrm{rank}\begin{bmatrix}\tilde{\boldsymbol{R}}_A \\ \tilde{\boldsymbol{R}}_C\end{bmatrix} \tag{26}$$

Hence, for our purposes, we can compute the rows of $\tilde{\boldsymbol{R}}_C$ instead of those of $\tilde{\boldsymbol{R}}_B$. Compared to (17), here the algebra required is *really* tedious and, due to space limitations, we directly provide the final results. For $\boldsymbol{f}_1=\boldsymbol{g}_{v_{i_k}}$ we have:

$$\bar{\nabla}_{\boldsymbol{p}_i}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm} = -\left[\frac{1}{d_{ij}^2}\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\boldsymbol{R}_i^T\boldsymbol{R}_i\boldsymbol{e}_k\boldsymbol{R}_i\boldsymbol{\beta}_{ij}\right]^T \tag{27}$$

$$\bar{\nabla}_{\boldsymbol{p}_j}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm} = \left[\frac{1}{d_{ij}^2}\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\boldsymbol{R}_i^T\boldsymbol{R}_i\boldsymbol{e}_k\boldsymbol{R}_i\boldsymbol{\beta}_{ij}\right]^T \tag{28}$$

$$\bar{\nabla}_{\boldsymbol{R}_i}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm} = \boldsymbol{0}^T, \qquad \bar{\nabla}_{\boldsymbol{R}_j}L^1_{\boldsymbol{g}_{v_{i_k}}}\beta_{ijm} = \boldsymbol{0}^T \tag{29}$$

The expressions in (29) are zero due to a numerical cancellation inside the skew operator and the fact that $L^1_{\boldsymbol{g}_{\boldsymbol{v}_{i_k}}}\beta_{ijm}$ does not depend on $\boldsymbol{R}_j$. For $\boldsymbol{f}_1=\boldsymbol{g}_{v_{j_k}}$ we have:

$$\bar{\nabla}_{\boldsymbol{p}_i}L^1_{\boldsymbol{g}_{v_{j_k}}}\beta_{ijm} = \left[\frac{1}{d_{ij}^2}\boldsymbol{R}_i\boldsymbol{\beta}_{ij}\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\boldsymbol{R}_i^T\boldsymbol{R}_j\boldsymbol{e}_k\right]^T \tag{30}$$

$$\bar{\nabla}_{\boldsymbol{p}_j}L^1_{\boldsymbol{g}_{v_{j_k}}}\beta_{ijm} = -\left[\frac{1}{d_{ij}^2}\boldsymbol{R}_i\boldsymbol{\beta}_{ij}\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\boldsymbol{R}_i^T\boldsymbol{R}_j\boldsymbol{e}_k\right]^T \tag{31}$$

$$\bar{\nabla}_{\boldsymbol{R}_i}L^1_{\boldsymbol{g}_{v_{j_k}}}\beta_{ijm} = \left[\frac{1}{d_{ij}}\mathrm{skew}\left(2\boldsymbol{R}_i^T\boldsymbol{R}_j\boldsymbol{e}_k\boldsymbol{e}_m^T\boldsymbol{P}_{ij}\right)^\vee\right]^T \tag{32}$$

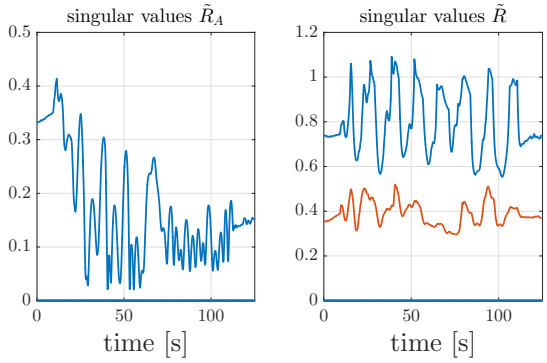


Fig. 1: Behavior of the last 8 singular values of the matrix $\tilde{\boldsymbol{R}}_A$ (left) and of the matrix $\tilde{\boldsymbol{R}}$ (right)

$$\bar{\nabla}_{\boldsymbol{R}_j} L^1_{\boldsymbol{g}_{v_{j_k}}} \beta_{ijm} = \left[\frac{1}{d_{ij}} \operatorname{skew}\left(2\boldsymbol{R}_j^T \boldsymbol{R}_i \boldsymbol{P}_{ij} \boldsymbol{e}_m \boldsymbol{e}_k^T \right)^\vee \right]^T \tag{33}$$

Regarding the case $\boldsymbol{f}_1 = \boldsymbol{g}_{w_{i_k}}$, the Lie derivative $L^1_{\boldsymbol{g}_{w_{i_k}}} \beta_{ijm}$ depends on the states only through $\boldsymbol{\beta}_{ij}$. Since we have separated the contribution of $\dot{\boldsymbol{\beta}}_{ij}$ in (24), we have that $\bar{\nabla}_{\boldsymbol{q}} L^1_{\boldsymbol{g}_{w_{i_k}}} \beta_{ijm} = 0$. Finally, as already mentioned, all the other first order Lie derivatives are zero, so they cannot contribute to the rank of $\tilde{\boldsymbol{R}}$.

C. Numerical verification of the ranks of $\tilde{\boldsymbol{R}}_A$ and $\tilde{\boldsymbol{R}}$

Sects. III-A and III-B provided a complete analytical expression of the different terms of the matrix $\tilde{\boldsymbol{R}}$. Notice that these matrices (and their ranks) depend only on the position and rotations of the agents, on their velocities (i.e., inputs).

After building the matrix (25) we found that, for a formation of 3 agents in random positions, the rank of the $\tilde{\boldsymbol{R}}_A$ is equal to 11 $(6N-7)$ while the one of the $\tilde{\boldsymbol{R}}$ to 12 $(6N-6)$. As an example, we plot in Fig. 1 the last 8 singular values of the two matrices as a function of time along the trajectories of the experiment discussed in Sect. V. Note that, as with any dynamic observability result, this result does not imply that the global scale can be recovered under *any arbitrary* choice of inputs. Instead, it is necessary to choose inputs that "excite" this mode (intuitively, where the agents move approximately perpendicularly to the bearing). This is in line with what was found in [21].

Note that at the beginning of the paper we claimed that this algorithm is bearing-only while in the previous terms, almost everywhere, it appears the distance d_{ij}. This quantity is retrievable through the EKF which is the topic of next Sect. IV

IV. A Multi-Agent Extended Kalman Filter

This section describes the design of an Extended Kalman Filter (EKF) on Lie groups to empirically verify the ideas of the previous sections. In particular, the EKF will provide an estimate $\hat{\boldsymbol{q}}_i = (\hat{\boldsymbol{p}}_i, \hat{\boldsymbol{R}}_i)$ of the configuration $\boldsymbol{q}_i$ of each agent. There are, however, two issues that need to be addressed. First, as already seen, the state of the system can be estimated only up to a global rotation and translation. We fix this ambiguity by choosing a moving reference frame that moves with the first agent. Of course, the uncertainty associated with this agent will always be zero by construction. A more representative choice would probably be to consider an "average" reference frame placed at the centroid of the formation; this, however, would significantly complicate the derivation of the filter, and it is therefore outside the scope of this paper. Second, actual implementations of any filter must work in discrete time. We therefore need to discretize our dynamical model (1). As in [31], we use a simple 1-step Euler forward approach, which is equivalent to assuming constant input velocites $\boldsymbol{v}_i, \boldsymbol{w}_i$ between discretization instants. With this assumption, our system (1) becomes

$$\boldsymbol{q}_{i,k+1} = \begin{bmatrix} \boldsymbol{p}_{i,k} + dt\boldsymbol{v}_{i,k} \\ \exp_{\boldsymbol{R}_{i,k}}\left(dt\boldsymbol{R}_{i,k}(\boldsymbol{w}_{i,k})^\wedge \right) \end{bmatrix} = \boldsymbol{f}_i(\boldsymbol{q}_{i,k}, \boldsymbol{u}_{i,k}), \tag{34}$$

where $u_{i,k} = \operatorname{stack}(v_{i,k}, w_{i,k})$, dt is the length of time of the discretization interval, and $\exp_{\boldsymbol{R}}(\cdot)$ is the exponential map at $\boldsymbol{R} \in \mathrm{SO}(3)$; the exponential map can be computed as $\exp_{\boldsymbol{R}}(\boldsymbol{R}\hat{\boldsymbol{w}}) = \boldsymbol{R}\exp_{\boldsymbol{I}}(\hat{\boldsymbol{w}})$, where the exponential at the identity $\exp_I$, which corresponds to the matrix exponential, can be computed using the Rodriguez's formula [2]. Moreover, we define an operator $\exp_{\boldsymbol{q}_k}(\tau_k)$ with $\boldsymbol{q}_k = (\boldsymbol{p}_k, \boldsymbol{R}_k) \in \mathrm{SE}(3)$ and $\tau_{\mathbf{k}} = (\boldsymbol{v}_k, \boldsymbol{w}_k) \in T_{\boldsymbol{q}}\mathrm{SE}(3)$. This operator for the position part of $\boldsymbol{q}_k$ corresponds to the operation $\boldsymbol{q}_k + \boldsymbol{v}_k dt$ and for the rotation part of $\boldsymbol{q}_k$ corresponds to the usual $\exp_{\boldsymbol{R}_k}(\hat{\boldsymbol{w}}_k)$ (this notation simplifies the filter equations). Note that (34) is used for $i > 1$; for agent $i = 1$ we have $\boldsymbol{q}_{i,k} = \boldsymbol{0}$, given our choice of the reference frame. Note also that the output function $\boldsymbol{h}_{ij}$ in (3), corresponding to the edge (i, j), should be redefined to explicitly take into account the measurement noise (denoted with $\boldsymbol{o}_{ij}$)

$$\tilde{\boldsymbol{h}}_{ij}(\boldsymbol{q}_i, \boldsymbol{q}_j, \boldsymbol{o}_{ij}) = \boldsymbol{R}_i^{\mathrm{T}} \frac{\boldsymbol{p}_j - \boldsymbol{p}_i + \boldsymbol{o}_{ij}}{\|\boldsymbol{p}_j - \boldsymbol{p}_i + \boldsymbol{o}_{ij}\|} \tag{35}$$

Similarly, we define $\tilde{\boldsymbol{f}}_i(\boldsymbol{q}_i, \boldsymbol{u}_i, \boldsymbol{n}_i) = \boldsymbol{f}(\boldsymbol{q}_i, \boldsymbol{u}_i + \boldsymbol{n}_i)$, where $\boldsymbol{n}_i \in T_{\boldsymbol{q}_i} SE(3)$ is the noise.

The overall system for the entire network is then

$$\boldsymbol{q}_k = \tilde{\boldsymbol{f}}(\boldsymbol{q}_{k-1}, \boldsymbol{u}_k, \boldsymbol{n}_k), \quad \boldsymbol{\beta}_{ij} = \tilde{\boldsymbol{h}}(\boldsymbol{q}_j, \boldsymbol{q}_j, \boldsymbol{o}_k), \tag{36}$$

with $\boldsymbol{u}_k = \operatorname{stack}(\{u_{i,k}\})$, and where we added process and measurement noises $\boldsymbol{n}_k, \boldsymbol{o}_k$ ($\boldsymbol{o}_k = \operatorname{stack}(\{\boldsymbol{o}_{ij}\})$ introduced in (35)), which are assumed to be zero mean multivariate Gaussian with covariance (block diagonal) matrixes $\boldsymbol{Q}_k, \boldsymbol{R}_k$. The EKF is based on the linearization of the system (36):

$$\boldsymbol{F}_k = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{q}}\right|_{\substack{\hat{\boldsymbol{x}}_{k-1|k-1}, \\ \boldsymbol{u}_k, \boldsymbol{n}_k = \boldsymbol{0}}}, \quad \boldsymbol{H}_k = \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{q}}\right|_{\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{o}_k = \boldsymbol{0}}, \tag{37}$$

$$\boldsymbol{L}_k = \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{n}}\right|_{\substack{\hat{\boldsymbol{x}}_{k-1|k-1}, \\ \boldsymbol{u}_{k-1}, \boldsymbol{n}_k = \boldsymbol{0}}}, \quad \boldsymbol{M}_k = \left.\frac{\partial \boldsymbol{h}}{\partial \boldsymbol{o}}\right|_{\hat{\boldsymbol{x}}_{k|k-1}, \boldsymbol{o}_k = \boldsymbol{0}}. \tag{38}$$

In our case, the matrix $\boldsymbol{F}$ is block diagonal with blocks given by

$$\boldsymbol{F}_i \dot{\boldsymbol{q}}_i = \frac{\partial \boldsymbol{f}_i}{\partial \boldsymbol{q}_i} \dot{\boldsymbol{q}}_i = \begin{pmatrix} \dot{\boldsymbol{p}}_i + dt\dot{\boldsymbol{R}}_i \boldsymbol{v}_i + dt\boldsymbol{R}_i \dot{\boldsymbol{v}}_i \\ D_{\boldsymbol{R}_i} \dot{\boldsymbol{R}}_i^\vee + dt D_{\boldsymbol{w}_i} \dot{\boldsymbol{w}}_i \end{pmatrix}. \tag{39}$$

$D_{\boldsymbol{R}_i} = \mathrm{D}_{\boldsymbol{R}_i} \exp_{\boldsymbol{R}_i}(dt\boldsymbol{R}_i\hat{\boldsymbol{w}}_i)$ and $D_{\boldsymbol{w}_i} = \mathrm{D}_{\boldsymbol{w}_i} \exp_{\boldsymbol{R}_i}(dt\boldsymbol{w}_i)$ denote the differentials of the exponential map with respect to $\boldsymbol{R}$ and $\boldsymbol{w}$, respectively; details on the implementation of

these differential can be found in [43]. The matrix $\boldsymbol{H}_k$ is actually the same as the matrix $\tilde{\boldsymbol{R}}_A$ computed in Sect. III-A. The matrices $\boldsymbol{L}_k$ and $\boldsymbol{M}_k$ can be computed similarly to $\boldsymbol{F}_k$ and H_k.

The *prediction* step of the EKF is given by the state estimate

$$\hat{\boldsymbol{q}}_{k|k-1} = \boldsymbol{f}(\hat{\boldsymbol{q}}_{k-1|k-1}, \boldsymbol{u}_{k-1}) \tag{40}$$

and the covariance matrix estimate

$$\boldsymbol{P}_{k|k-1} = \boldsymbol{F}_{k-1}\boldsymbol{P}_{k-1|k-1}\boldsymbol{F}_{k-1}^T + \boldsymbol{L}_{k-1}\boldsymbol{Q}_{k-1}\boldsymbol{L}_{k-1}^T. \tag{41}$$

The *update* step of the filter consists of the state estimate

$$\begin{aligned}\hat{\boldsymbol{q}}_{k|k} = \exp_{\hat{\boldsymbol{q}}_{k|k-1}}\Big(&\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T + \\ &+ \underbrace{\left(\boldsymbol{H}_k\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{M}_k\boldsymbol{R}_k\boldsymbol{M}_k^T\right)}_{\boldsymbol{S}_k}{}^{-1}\Big)\tilde{\boldsymbol{y}}_k\end{aligned} \tag{42}$$

where $\tilde{\boldsymbol{y}}_k$ is the difference between the measured and estimated bearings.

Moreover, the covariance matrix estimate is given by

$$\boldsymbol{P}_{k|k} = \left(\boldsymbol{I} - \left(\boldsymbol{P}_{k|k-1}\boldsymbol{H}_k^T + \boldsymbol{S}_k^{-1}\right)\boldsymbol{H}_k\right)\boldsymbol{P}_{k|k-1} \tag{43}$$

Sect. V below shows a result of our (centralized) EKF.

V. Experimental Results

In order to validate the presented ideas, extensive simulations in a Matlab environment have been performed considering cases with both process and measurement noise. In addition to simulations also experiments have been performed with real quadrotor UAVs[3]. In this section an experiment with 3 agents (2 real quadrotor UAVs and a fixed *virtual* agent) will be presented. The experiments were performed in our flying arena which has a flying volume of 6.5 m x 6.5 m x 3 m and is equipped with a Vicon motion capture system. This was used for obtaining the ground truth and for reconstructing the body-frame bearing measurements $\boldsymbol{\beta}_{ij}$ that would have been obtained by the onboard cameras.

For the experiment presented in this section 2 MK-Quadro by Mikrokopter have been used. The usual Mk-Quadro was extended with an ODROID-XU4 Linux computer running ROS and a set of packages based on GenoM3 [44] for implementing the low-level flight control. The high level controller was implemented in Matlab. The third *virtual* agent of the formation which fix the ambiguity discussed in Sect. IV has been placed, for convenience, in the origin of the motion capture system reference frame. This agent will be the one w.r.t. the other 2 UAVs will refer their measurements.

During the experiment reported in Figs. 2 and 3 the agent 2 was moving mainly back and forth on the x axis while the agent 3 was following an ellipsoidal trajectory. An user was also able to give velocity inputs to the UAVs through a joystick. The trajectories described above allows to highlight the fact that, in order to collectively estimate the scale, the agents need to move along *exciting* trajectories [21]. Indeed, from Figs. 2a and 2b it is possible to notice that the variance associated to the coordinates which are only *ligthly excited* (y, z coordinates on the agent 2 and 3) remains higher than the ones which are *strongly excited* (x coordinates on the agents 2 and 3). Fig. 2c shows the behavior of the true and estimated distance between the agents 2 and 3 and it is clear that the EKF, even starting far from the true values (both for position and orientations (Fig. 3)) it is able to recover the true distance d_{23}, hence the scale of the formation.

The orientation error showed in Fig. 3 is defined as:

$$\boldsymbol{R}_{err_i} = \frac{1}{2}\operatorname{tr}\left(\boldsymbol{I} - \hat{\boldsymbol{R}}_i^T\boldsymbol{R}_i\right),\ \forall i \in \mathcal{V} \tag{44}$$

VI. Conclusions and future works

In this paper we considered the problem of scale estimation in localization with bearing-only measurements and known agent velocities. By applying nonlinear observability theory and Riemannian geometry, we extended existing results on the theory of rigidity and introduced the notion of *Dynamic Bearing Observability Matrix* (DBOM). Using experiments with 2 quadrotor UAVs, we have shown that the global scale is indeed observable, and that it can be recovered by employing a centralized EKF on $\mathrm{SE}(3)$.

The preliminary results of this work open several interesting future research direction, such as *1)* providing a full analytical characterization of the nullspace of the DBOM and critical agent configurations, *2)* verifying if the inclusion of higher order Lie derivatives can provide additional insights, *3)* designing input signals that minimize the uncertainty in localization (using [21] as inspiration), *4)* studying distributed implementations of the EKF (building upon, e.g., [34], [45]), *5)* investigating the use of real vision-based measurements (with, e.g., AprilTag [46]), *6)* applying similar ideas to different types of agents (e.g. second-order integrators or unicycle dynamics) and measurements (e.g., distance-only) and, finally, *7)* extending the proposed ideas to the case of multi-agent systems with unknown input as in [47].

[3]Video of the experiments: `https://goo.gl/izyYk3`

References

[1] K. Sreenath, T. Lee, and V. Kumar, "Geometric control and differential flatness of a quadrotor uav with a cable-suspended load," in *IEEE Intl. Conf. on Dec. and Cont.* IEEE, 2013, pp. 2269–2274.

[2] Y. Ma, *An invitation to 3-D vision: from images to geometric models*. Springer, 2004.

[3] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. Cambridge University Press, 2004.

[4] E. S. Jones and S. Soatto, "Visual-inertial navigation, mapping and localization: A scalable real-time causal approach," *International Journal of Computer Vision*, vol. 30, no. 4, pp. 407–430, 2011.

[5] J. Aspnes, W. Whiteley, and Y. R. Yang, "A theory of network localization," vol. 5, no. 12, pp. 1663–1678, 2006.

[6] T. Eren, W. Whiteley, and P. N. Belhumeur, "Using angle of arrival (bearing) information in network localization," in *IEEE Intl. Conf. on Dec. and Cont.* IEEE, 2006, pp. 4676–4681.

[7] B. Shirmohammadi and C. Taylor, "Self localizing smart camera networks," *ACM Transactions on Sensor Networks*, 2010.

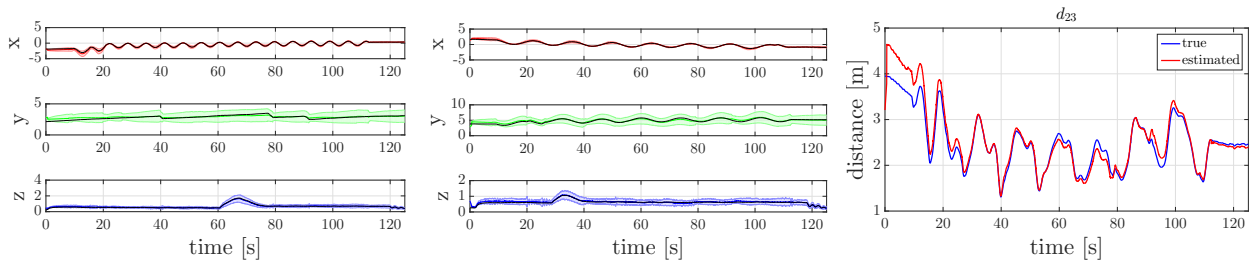
[8] A. N. Bishop, I. Shames, and B. Anderson, "Stabilization of rigid formations with direction-only constraints," in *IEEE Intl. Conf. on Dec. and Cont.*, 2011, pp. 746–752.

[9] A. Franchi and P. R. Giordano, "Decentralized control of parallel rigid formations with direction constraints and bearing measurements," in *IEEE Intl. Conf. on Dec. and Cont.*, 2012, pp. 5310–5317.

[10] S. Zhao, F. Lin, K. Peng, B. M. Chen, and T. H. Lee, "Distributed control of angle-constrained cyclic formations using bearing-only measurements," *Systems and Control Letters*, vol. 63, pp. 12–24, 2014.

[11] M. Brand, M. Antone, and S. Teller, "Spectral solution of large-scale extrinsic camera calibration as a graph embedding problem," in *IEEE European Conf. on Comp. Vision*, 2004, pp. 262–273.

[12] H. Li, "Multi-view structure computation without explicitly estimating motion," in *IEEE Conf. on Comp. Vision and Pat. Rec.*, 2010, pp. 2777–2784.



(a) (b) (c)

Fig. 2: Fig. 2a shows the behavior of the real position (x, y, z) (in meters) of the agent 2, in black, and the estimated one with the associated covariance respectively in red (x), green (y) and blue (z); Fig. 2b is the dual of Fig. 2a for the agent 3; Fig. 2c shows the behavior of the true (blue) and estimated (red) distance d_{23} between the agents 2 and 3.

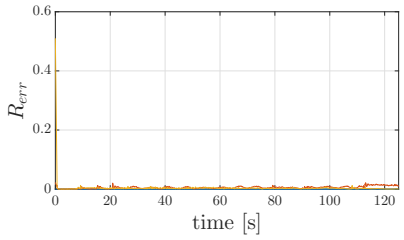


Fig. 3: Behavior of the orientation error introduced in (44)

[13] B. Servatius and W. Whiteley, "Constraining plane configurations in computer-aided design: Combinatorics of directions and lengths," *SIAM Journal on Discrete Mathematics*, vol. 12, no. 1, pp. 136–153, 1999.

[14] B. D. Anderson, C. Yu, J. M. Hendrickx *et al.*, "Rigid graph control architectures for autonomous formations," *IEEE Cont. Sys.*, vol. 28, no. 6, 2008.

[15] R. Tron, L. Carlone, F. Dellaert, and K. Daniilidis, "Rigid components identification and rigidity enforcement in bearing-only localization using the graph cycle basis," in *IEEE Am. Cont. Conf.*, Chicago, IL, Jul. 2015, pp. 3911–3918.

[16] D. Zelazo, P. Robuffo Giordano, and A. Franchi, "Bearing-Only Formation Control Using an SE(2) Rigidity Theory," in *2015 IEEE CDC*, Osaka, Japan, Dec. 2015, pp. 6121–6126.

[17] F. Schiano, A. Franchi, D. Zelazo, and P. Robuffo Giordano, "A rigidity-based decentralized bearing formation controller for groups of quadrotor UAVs," in *IEEE Intl. Conf. on Intell. Rob. and Sys.*, 2016.

[18] F. Schiano and P. Robuffo Giordano, "Bearing Rigidity Maintenance for Formations of Quadrotor UAVs," in *IEEE Intl. Conf. on Rob. and Aut.*, 2017.

[19] R. Tron, L. Carlone, F. Dellaert, and K. Daniilidis, "Rigid components identification and rigidity enforcement in bearing-only localization using the graph cycle basis," in *IEEE Am. Cont. Conf.*, 2015.

[20] G. Michieletto, A. Cenedese, and A. Franchi, "Bearing rigidity theory in $SE(3)$," in *IEEE Intl. Conf. on Dec. and Cont.* IEEE, 2016, pp. 5950–5955.

[21] R. Spica and P. Robuffo Giordano, "Active Decentralized Scale Estimation for Bearing Based Localization," in *IEEE Intl. Conf. on Intell. Rob. and Sys.*, 2016.

[22] A. Martinelli, "Vision and imu data fusion: Closed-form solutions for attitude, speed, absolute scale, and bias determination," *IEEE Trans. Automat. Contr.*, vol. 28, no. 1, pp. 44–60, 2012.

[23] R. K. Williams and G. S. Sukhatme, "Observability in topology-constrained multi-robot target tracking," in *IEEE Intl. Conf. on Rob. and Aut.* IEEE, 2015, pp. 1795–1801.

[24] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Trans. Automat. Contr.*, vol. 22, no. 5, pp. 728–740, 1977.

[25] H. K. Khalil, *Nonlinear Systems*. Prentice-Hall, New Jersey, 1996.

[26] A. Martinelli and R. Siegwart, "Observability analysis for mobile robot localization," in *IEEE Intl. Conf. on Intell. Rob. and Sys.* IEEE, 2005, pp. 1471–1476.

[27] A. Isidori, *Nonlinear control systems*. Springer Science & Business Media, 2013.

[28] F. Bullo and A. D. Lewis, *Geometric control of mechanical systems: modeling, analysis, and design for simple mechanical control systems*. Springer Science & Business Media, 2004, vol. 49.

[29] F. L. Markley and J. L. Crassidis, *Fundamentals of spacecraft attitude determination and control*. Springer, 2014, vol. 33.

[30] T. Lee, M. Leok, and N. H. McClamroch, "Nonlinear robust tracking control of a quadrotor UAV on SE(3)," *Asian Journ. of Cont.*, vol. 15, no. 2, pp. 391–408, 2013.

[31] G. Loianno, M. Watterson, and V. Kumar, "Visual inertial odometry for quadrotors on SE(3)," in *IEEE Intl. Conf. on Rob. and Aut.* IEEE, 2016, pp. 1544–1551.

[32] R. Tron and R. Vidal, "Distributed 3-D localization of camera sensor networks from 2-D image measurements," *IEEE Trans. Automat. Contr.*, 2014.

[33] S. I. Roumeliotis and I. M. Rekleitis, "Propagation of uncertainty in cooperative multirobot localization: Analysis and experimental results," *Autonomous Robots*, vol. 17, no. 1, pp. 41–54, 2004.

[34] L. Luft, T. Schubert, S. I. Roumeliotis, and W. Burgard, "Recursive Decentralized Collaborative Localization for Sparsely Communicating Robots," *Robotics: Science and Systems*, 2016.

[35] T. Lee, "Geometric control of multiple quadrotor uavs transporting a cable-suspended rigid body," in *IEEE Intl. Conf. on Dec. and Cont.* IEEE, 2014, pp. 6155–6160.

[36] A. Barrau and S. Bonnabel, "The invariant extended kalman filter as a stable observer," *IEEE Trans. Automat. Contr.*, vol. 62, no. 4, pp. 1797–1812, 2017.

[37] D. Zelazo, A. Franchi, and P. Robuffo Giordano, "Rigidity Theory in SE(2) for Unscaled Position Estimation using only Bearing Measurements," in *2014 ECC*, Strasbourg, France, Jun. 2014, pp. 2703–2708.

[38] M. P. do Carmo, *Riemannian geometry*. Boston, MA: Birkhäuser, 1992.

[39] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 2012.

[40] Y. Ma, J. Kosecka, and S. Sastry, "Optimization criteria and geometric algorithms for motion and structure estimation," *International Journal of Computer Vision*, vol. 44, no. 3, pp. 219–249, 2001.

[41] T. Eren, W. Whiteley, A. S. Morse, P. N. Belhumeur, and B. D. Anderson, "Sensor and network topologies of formations with direction, bearing, and angle information between agents," in *IEEE Intl. Conf. on Dec. and Cont.*, vol. 3. IEEE, 2003, pp. 3064–3069.

[42] F. Schiano and R. Tron, "The Dynamic Bearing Observability Matrix Nonlinear Observability and Estimation for Multi-Agent Systems," Tech. Rep., accessed: 2018-02-20. [Online]. Available: http://www.irisa.fr/lagadic/pdf/2018_icra_schiano_technical_report.pdf

[43] R. Tron, "Distributed optimization on manifolds for consensus algorithms and camera network localization," Ph.D. dissertation, The Johns Hopkins University, 2012.

[44] A. Mallet, C. Pasteur, M. Herrb, S. Lemaignan, and F. Ingrand, "GenoM3: Building middleware-independent robotic components," in *IEEE Intl. Conf. on Rob. and Aut.* IEEE, 2010, pp. 4627–4632.

[45] Z. Lendek, R. Babuska, and B. D. Schutter, "Distributed Kalman filtering for multiagent systems," *IEEE Eu. Cont. Conf.*, vol. 19, pp. 2193–2200, 2007.

[46] J. Wang and E. Olson, "Apriltag 2: Efficient and robust fiducial detection," in *IEEE Intl. Conf. on Intell. Rob. and Sys.* IEEE, 2016, pp. 4193–4198.

[47] A. Martinelli, "Nonlinear Unknown Input Observability: Extension of the Observability Rank Condition," *IEEE Transactions on Automatic Control*, vol. 9286, pp. 1–16, 2018.