Intrinsic Consensus on $SO(3)$ with Almost-Global Convergence

Roberto Tron, Bijan Afsari and René Vidal

***Abstract*— In this paper we propose a discrete time protocol to align the states of a network of agents evolving in the space of rotations $SO(3)$. The starting point of our work is Riemannian consensus, a general and intrinsic extension of classical consensus algorithms to Riemannian manifolds. Unfortunately, this algorithm is guaranteed to align the states only when the initial states are not too far apart. We show how to modify Riemannian consensus so that the states of the agents can be aligned, in practice, from almost any initial condition. While we focus on the specific case of $SO(3)$, we hope that this work will represent the first step toward more general results.**

I. INTRODUCTION

Assume a network of N agents, where the state of each agent is represented by a 3-D rotation matrix and each agent can communicate with a small set of neighbors. In this paper, we consider the problem of finding a discrete time protocol that the agents can use to asymptotically synchronize their states. This and similar problems appear in a variety of situations such as vehicle coordination [1], attitude coordination [2], [3], pose averaging [4], [5] or even camera network localization from images [6].

Natural candidates for solving this problem are *consensus algorithms* [7], where the state of each node is updated using a combination of the states from the neighbors. An attractive property of consensus algorithms is that they are completely autonomous and do not rely on any central coordination.

Classical consensus algorithms assume that the states of the agents evolve in Euclidean space ($\mathbb{R}^d$). However, in this paper we are interested in extensions of consensus to Riemannian manifolds. Existing work in this area can be divided in two categories. The first category comprises *extrinsic* algorithms [8], [9], [10], [4]. These solutions rely on the embedding of the manifold (*e.g.*, $SO(3)$) in an Euclidean ambient space (*e.g.*, $\mathbb{R}^{3\times 3}$), where the usual consensus algorithms can be employed. The states are then obtained through projections. The downside of these algorithms is that they rely on a specific choice of the embedding and projection operations, and that they cannot be easily generalized. On the other hand, these algorithm can show almost-global convergence.

The second category comprises *intrinsic* algorithms, such as [5], [11], [12]. These algorithms are formulated as a distributed minimization problem and depend only on the intrinsic geometric properties of the manifold, such as the definition of a metric and of geodesics. However, due to the geometry of the manifold, the underlying optimization problem is typically not convex. This results in algorithms that, depending on the network topology and on the initial conditions, might get trapped near undesired configurations where the states are not aligned (as we will see in §III).

Comparing the two categories, a natural question arises: is it possible to achieve almost-global convergence with an intrinsic formulation? In this paper we give a partial affirmative answer by proposing an intrinsic consensus algorithm with almost-global convergence on $SO(3)$.

Related work. This paper builds on an existing almost-global convergent consensus algorithm on the circle [13]. We extend the approach to $SO(3)$, and we pass from a continuous to a discrete time formulation. Other works treat the state alignment problem with a state-feedback approach, where the control inputs for the agents are designed. Proofs of convergence then use either Lyapunov stability ([14], [15]) or passivity ([16], [3] and references therein). However, all these works consider a continuous timeformulation, and they provide only local convergence results.

An expanded version of this article is available in [17].

II. NOTATION AND PRELIMINARIES

We model the network of N agents using an undirected connected graph $G = (V, E)$. The vertices $V = \{1, \dots, N\}$ represent the nodes of the network while the edges $E \subseteq V \times V$ represent the communication links. The set of neighbors of node i is denoted as $N_i = \{j \in V | \ (i, j) \in E\}$. We indicate the *maximum degree* of the graph G as $\operatorname{Deg}(G) = \max_{i \in V}\{|N_i|\}$, where $|N_i|$ is the number of neighbors of node i. We denote as $\operatorname{Diam}(G)$ the diameter of the graph G, *i.e.*, the maximum length of the shortest path between any two vertices in the graph [18].

We associate to each agent $i \in V$ a state $R_i \in SO(3)$, where $SO(3) = \{R \in \mathbb{R}^{3\times 3} : R^T R = I, \det(R) = 1\}$ is the space of three-dimensional rotations. The *tangent space* of $SO(3)$ at a rotation R given by $T_R SO(3) = \{RV : V \in \mathfrak{so}(3)\}$, where $\mathfrak{so}(3)$ is the space of 3×3, skew symmetric matrices. We can identify a tangent vector $W \in T_R SO(3)$ with a vector $w \in \mathbb{R}^3$ using the usual *hat* $(\cdot)^\wedge$ and *vee* $(\cdot)^\vee$ operators, given by the relations

$$w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \underset{(\cdot)^\vee}{\overset{(\cdot)^\wedge}{\rightleftarrows}} W = R \begin{bmatrix} 0 & -w_3 & w_2 \\ w_3 & 0 & -w_1 \\ -w_2 & w_1 & 0 \end{bmatrix}. \quad (1)$$

With the identification in (1), given two tangent vectors $W_1, W_2 \in T_R SO(3)$ and the corresponding vector representations $w_1, w_2 \in \mathbb{R}^3$, the standard metric for $SO(3)$ is given by:

$$\langle W_1, W_2 \rangle = \langle \hat{w}_1, \hat{w}_2 \rangle = w_1^T w_2. \quad (2)$$

This work was supported by JHU startup funds and by grants NSF CNS-0834470, NSF ECCS-0941463 and ONR N000140910084.

The authors are with the Center for Imaging Science, Johns Hopkins University, Baltimore MD, 21202, USA

For a given rotation $R \in SO(3)$, the *exponential* and *logarithm* map are denoted, respectively, as $\exp_R : T_R SO(3) \to SO(3)$ and $\log_R : U_R \to T_R SO(3)$, where $U_R \subset SO(3)$ is the maximal set containing R for which $\exp_R$ is diffeomorphic. For convenience, we also define $\mathrm{Log} : U_I \to \mathbb{R}^3$, the vectorized version of the logarithm map at the identity:

$$\mathrm{Log}(R) = (\log_I(R))^\vee \in \mathbb{R}^3, \tag{3}$$

where $R \in SO(3)$. See [17] for closed form expressions. Note that $\mathrm{Log}(R^T S) = (R^T \log_R S)^\vee$ for $R, S \in SO(3)$ [5].

Also, for any edge $(i,j) \in E$ we use the shorthand notation

$$\theta_{ij} = d(R_i, R_j) \quad \text{and} \quad u_{ij} = \frac{\mathrm{Log}(R_i^T R_j)}{\|\mathrm{Log}(R_i^T R_j)\|} \tag{4}$$

to indicate, respectively, the geodesic distance in $SO(3)$ between the two rotations R_i and R_j and the normalized rotation axis of $R_i^T R_j$. Unless necessary, we omit the explicit dependence of θ_{ij} and u_{ij} on their arguments.

Note the following equivalences:

$$\|\mathrm{Log}(R_i^T R_j)\| = \|\log_{R_i}(R_j)\| = d(I, R_i^T R_j) = \theta_{ij}. \tag{5}$$

For any given rotation $R \in SO(3)$, we denote as $\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R)$ the matrix representation of the differential of the normalized logarithm. More precisely, let $R(t)$ be a smooth curve in $SO(3)$ such that $R(0) = R_0$ and $\dot{R}(0) = W$, then

$$\frac{\mathrm{d}}{\mathrm{d}t} \frac{\mathrm{Log}(R)}{\|\mathrm{Log}(R)\|}\bigg|_{t=0} = \mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R_0) W^\vee. \tag{6}$$

This matrix and its spectral decomposition can be computed explicitly (see [17] for the proof):

Proposition 1: Given $R \in SO(3)$, let $\theta = \|Log(R)\|$ and $u = \frac{\mathrm{Log}(R)}{\|\mathrm{Log}(R)\|}$. The matrix representations in (6) is given by $\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R) = \frac{1}{2}\big(\hat{u} - \cot(\frac{\theta}{2})\hat{u}^2\big)$. One of the eigenvalues is zero with eigenvector u. The other eigenvalues are $\alpha(\theta) \pm \mathrm{j}\beta(\theta)$, where $\alpha(\theta) = \frac{\theta}{2}\cot(\frac{\theta}{2})$, $\beta(\theta) = \frac{\theta}{2}$ and $\mathrm{j} = \sqrt{-1}$ is the unit imaginary number. The corresponding complex conjugate eigenvectors $v, \bar{v} \in \mathbb{C}^3$ are orthogonal to u.

In the following, we use the product manifold $SO(3)^N$, which is the N-fold cartesian product of $SO(3)$ with itself. We use $\boldsymbol{R} = \{R_i\}_{i \in V}$ to indicate a point in $SO(3)^N$ and $\boldsymbol{W} = \{W_1, \ldots, W_N\}$, $W_i \in T_{R_i}SO(3)$ to indicate a tangent vector at $\boldsymbol{R}$. We use the natural metric $\langle \boldsymbol{V}, \boldsymbol{W} \rangle = \sum_{i=1}^N \langle V_i, W_i \rangle$. Geodesics, exponential maps, and gradients in $SO(3)$ can then be easily obtained by using the respective definitions on each copy of $SO(3)$. Here and in the following, we use boldface letters to denote N-tuples where each element represents a quantity related to one of the nodes.

Given a function $\varphi : SO(3)^N \to \mathbb{R}$ twice differentiable at a point $\boldsymbol{R}_0 \in SO(3)^N$, we denote the gradient of φ as $\mathrm{grad}\,\varphi(\boldsymbol{R}_0)$ and its Hessian as $\mathrm{Hess}\,\varphi(\boldsymbol{R}_0)$ [19]. By definition, we have the properties

$$\langle \mathrm{grad}\,\varphi(\boldsymbol{R}_0), \boldsymbol{W} \rangle = \frac{\mathrm{d}}{\mathrm{d}t}\varphi(\boldsymbol{R}(t))\Big|_{t=0}, \tag{7}$$

and

$$\langle \boldsymbol{W}, \mathrm{Hess}\,\varphi(\boldsymbol{R}_0)\boldsymbol{W} \rangle = \frac{\mathrm{d}^2}{\mathrm{d}t^2}\varphi(\boldsymbol{R}(t))\Big|_{t=0}, \tag{8}$$

where $R(t)$ is a smooth curve in $SO(3)^N$ such that $\boldsymbol{R}(0) = \boldsymbol{R}_0$ and $\dot{\boldsymbol{R}}(0) = \boldsymbol{W}$. The gradient of the distance from a fixed rotation $S \in SO(3)$ is given by

$$\mathrm{grad}_R\, d(S, R) = -\frac{\log_R(S)}{\|\log_R(S)\|}. \tag{9}$$

Note that at the minimum ($d(S,R) = 0$) and at the maximum ($d(S,R) = \pi$) of the distance, the function is continuous but not differentiable.

Given the twice differentiable function φ and an initial point $\boldsymbol{R}_0 \in SO(3)^N$, it is possible to define a steepest gradient descent algorithm on $SO(3)^N$ with constant step size $\varepsilon > 0$, as shown by Algorithm 1.

Algorithm 1 A Riemannian steepest gradient descent algorithm with fixed step size for differentiable functions

---

Input: An initial rotations $\boldsymbol{R}_0 \in SO(3)^N$, a step size $\varepsilon > 0$

1) **Initialize** $\boldsymbol{R}(0) = \boldsymbol{R}_0$
2) **For** $k \in \mathbb{N}$, **repeat**

$$\boldsymbol{W}(k) = -\,\mathrm{grad}\,\varphi(\boldsymbol{R}(k)) \tag{10}$$

$$\boldsymbol{R}(k+1) = \exp_{\boldsymbol{R}(k)}(\varepsilon \boldsymbol{W}(k)) \tag{11}$$

At each iteration k, the algorithm moves from the current state $x(k)$ to a new one along the geodesic in the direction opposite to the gradient. Under some conditions on the step size ε, $x(k)$ converges to the set of critical points of φ [20], [21]. Note that (9) is not defined where φ is non-differentiable. Therefore, (10) needs to be modified on an application-specific basis to handle these special cases, as we do in §III.

III. PROPOSED SOLUTION AND AN ILLUSTRATIVE EXAMPLE

Similarly to [11], we construct our proposed consensus algorithm by applying Algorithm 1 to the following function:

$$\varphi\big(\{R_i\}_{i \in V}\big) = \sum_{(i,j) \in E} f(\theta_{ij}), \tag{12}$$

where f is a reshaping function, as defined next.

Definition 2: We say that $f : [0, \pi] \to \mathbb{R}$ is a *reshaping function* if it satisfies the following assumptions:

- f is twice differentiable on $(0, \pi)$.
- $f(0) = 0$.
- $\dot{f}(0) = 0$, $\dot{f}(\theta) > 0$ for $\theta \in (0, \pi]$.

In the definition we used the notation $\dot{f}$ and $\ddot{f}$ to denote, respectively, the first and second derivative of f with respect to its only argument. We also use $\dot{f}^-$ and $\dot{f}^+$ to denote the left and right derivative of f, respectively.

Thanks to Def. 2, φ is at least twice differentiable except when $\theta_{ij} = \frac{\pi}{2}$ for at least one edge $(i,j) \in E$, in which case it is only continuous (see also §II).

In this paper, we propose the reshaping function

$$f(\theta) = a f_0(\theta) \text{ where } f_0(\theta) = \frac{1}{b} - \big(\frac{1}{b} + \theta\big)\exp(-b\theta), \tag{13}$$

$a = \frac{\pi^2}{2 f_0(\pi)}$ and $b \in \mathbb{R}$ is a "sufficiently large" constant. The reasons behind this specific choice of f and b will be clear

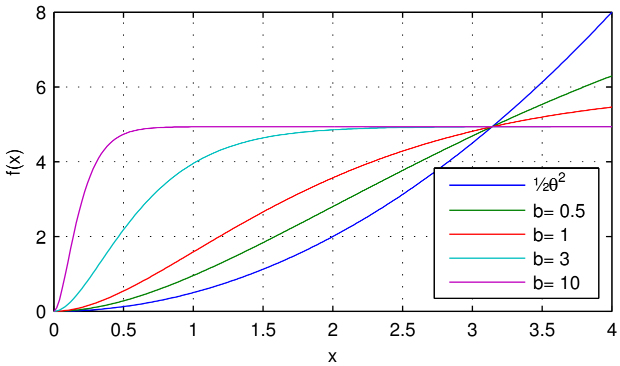
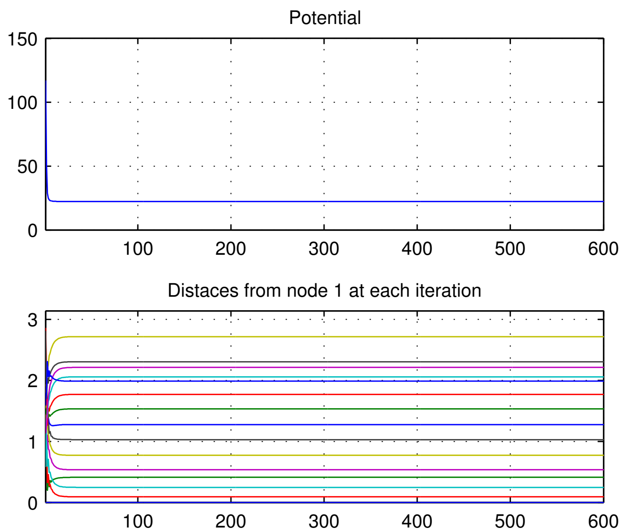
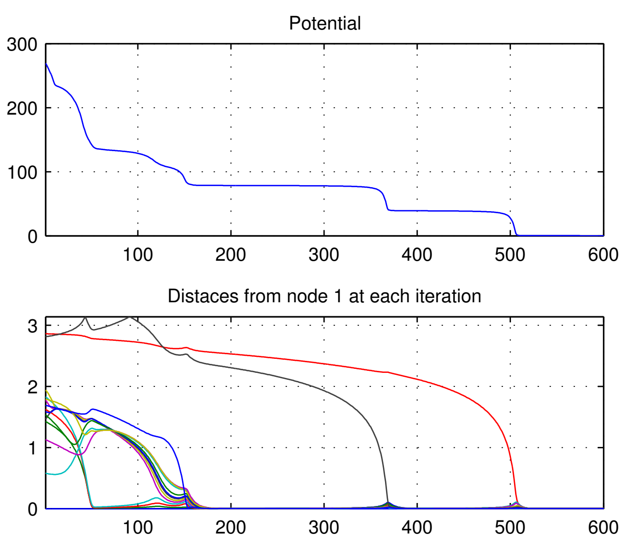


Fig. 1: Examples of the proposed reshaping function



(a)



(b)

Fig. 2: Example of consensus on a network of 10 nodes in a 4-regular graph with (a) the cost in [11] versus (b) the proposed reshaped cost ($b = 3$). The initial states are nearly uniformly distributed along a closed geodesic

by the end of the paper. For now, it is sufficient to notice that (13) tends to $f(\theta) = \frac{1}{2}\theta^2$ when b goes to zero (see Fig. 1). In this case, we recover the cost of [11] as a particular case of our framework.

With this cost function, (11) in Algorithm 1 becomes:

$$R_i(k+1) = \exp_{R_i(k)}\left(\varepsilon \sum_{j:(i,j)\in E} \dot{f}(\theta_{ij}) \frac{\log_{R_i(k)}(R_j(k))}{\|\log_{R_i(k)}(R_j(k))\|}\right). \tag{14}$$

Note that $\log_{R_i}(R_j)$ does not exist when $d(R_i, R_j) = \pi$. In this case, one can randomly perturb $\boldsymbol{R}(k)$ and compute (14) at that point. As mentioned before, the protocol in [11] is derived when b tends to zero.

To illustrate the intuition behind the proposed protocol, and to understand the role played by the reshaping function f, we consider a network with $N = 10$ nodes connected in a 4-regular graph topology. Let $\boldsymbol{R}_0$ be a configuration where the rotations are evenly distributed on a closed geodesic in $SO(3)$. We set the initial states in the network to a slightly perturbed version of $\boldsymbol{R}_0$. We first run the protocol (14) with the cost function of [11] ($f(\theta) = \frac{1}{2}\theta^2$) for 1000 iterations. Fig. 2a shows the distances of each state from the node $i = 1$ after every iteration. This algorithm gets trapped in a local minimizer of the cost function, and the states of the nodes do not deviate significantly from the initial configuration.

We then repeat the same experiment, but using instead the protocol (14) with the proposed reshaped cost and $b = 3$. The results are shown in Fig. 2b. In this case, the alignment of the nodes is reached after about 800 iterations.

This is not an isolated example, and it is indicative of the general behavior of the algorithms. To demonstrate this, we repeated 1000 times both experiments of Fig. 2, but with a different initial state $\boldsymbol{R}_0$ chosen uniformly at random. Table I reports the number of times the algorithms reached alignment among the nodes. The Riemannian consensus algorithm in [11] succeeded slightly less than half of the times, while the proposed algorithm with $b = 3$ aligned the states in all cases.

The different behavior of the two algorithms can be understood by considering the fact that configurations such as $\boldsymbol{R}_0$, where the states follow a closed geodesic, are undesired stationary points where $\operatorname{grad}_{\boldsymbol{R}} \varphi(\boldsymbol{R}_0) = 0$. Due to the topology of $SO(3)$, such kind of configurations are unavoidable and, for the cost in [11], they are stable local minimizers: The algorithm does not deviate too much from these points even in the presence of small perturbations. On the other hand, as we will show, for the proposed cost these undesired stationary points are unstable saddle points: When Algorithm 1 passes them, it slows down but it eventually deviates until it reaches the global minimizer. This is comfirmed by the flat regions in the potential of Fig. 2b.

In summary, while *global* convergence may not be attainable (due to undesired stationary points), we can still obtain *almost-global* convergence, in that the only stable stationary points are global minimizers.

Paper outline. The rest of the paper is devoted to showing that, when b is large enough, the proposed protocol converges to the sub-manifold of consensus configuration from almost

Riemannian consensus	Cost in [11]	Reshaped cost with $b = 3$
Number of correct alignments	485	1000

TABLE I: Number of successful alignments reached from 1000 random initializations for the same network as in Fig. 2

any initial configuration (Thm. 16). Our proof requires consideration of many different aspects of the cost function (12) and the protocol (14). In the first part (§IV–VI) we show that all the *desired equilibria* of the protocol (*i.e.*, the global minimizers of φ) are stable, while all the *undesired equilibria* (*i.e.*, any other equilibria) are unstable. We do this by using *θ-neighborhood subgraphs* (§IV) and their relation with the Hessian of φ (§V), which depends on the properties of the reshaping function f (§VI). In the second part (§VII) we focus on the choice of the step-size ε, which, together with the previous results, gives our main convergence theorem.

IV. DESIRED AND UNDESIRED EQUILIBRIA

Notice that (14) describes a dynamical system on $SO(3)^N$, whose equilibria we label as follows:

1) *Desired equilibria:* these are the global minimizers of φ, i.e., those points at which $\varphi = 0$, φ is differentiable, and $\operatorname{grad}\varphi = 0$.
2) *Undesired equilibria:* these are either stationary points of φ (e.g., local but not global minimizers) or points at which φ is non-differentiable and the sum in (14) is zero for every node $i \in V$.

Both the desired and undesired equilibria are critical points of φ by definition but not every critical point of φ is an equilibria of the system. Note that the desired equilibria are not isolated point and constitute a connected set. Our goal is to show that (under appropriate conditions) the *set* of desired equilibria is the only asymptotically stable equilibria set, while all the others are unstable (see Thm. 16).

The distinction between desired equilibria and undesired equilibria is intuitively easy. However, to continue our analysis, we need to build on these definitions. In this section, we first characterize a non-trivial subset $\mathcal{S} \subset SO(3)^N$ that contains only global minimizers and no other critical points. Then, we give a result which associates undesired equilibria, the set $\mathcal{S}$ and the connectivity of a *θ-neighborhood sub-graph*.

A non-trivial set with only global minimizers. We now extend basic results from [11] to our setting, where φ includes an arbitrary reshaping function and the manifold of interest is $SO(3)$. Define the diagonal space of $SO(3)^N$ as

$$\mathcal{D} = \left\{\boldsymbol{S} = \{S, \dots, S\} \in SO(3)^N : S \in SO(3)\right\}. \quad (15)$$

This space represents all the possible consensus configurations of the network, where all the nodes are aligned. It corresponds also to the space of desired equilibria (global minimizers), as proven by the following proposition.

Proposition 3: If G is connected, then $\boldsymbol{S} \in \mathcal{D}$ if and only if $\boldsymbol{S}$ is a global minimizer of φ.
The proof, with any reshaping function respecting Def. 2, follows from [11, Prop. 1] with trivial modifications.

We define the set $\mathcal{S} \subset SO(3)^N$ as

$$\mathcal{S} = \{\boldsymbol{R} \in SO(3)^N \colon \exists S \in SO(3) \text{ s.t. } \max_{i \in V} d(R_i, S) < \frac{\pi}{2}\}. \quad (16)$$

Intuitively, $\mathcal{S}$ is a tube in $SO(3)^N$ centered around the diagonal space $\mathcal{D}$ and having a "square" section. We can now extend [11, Theorem 1] to the use of reshaping functions.

Theorem 4: A point $\boldsymbol{R}_0 \in \mathcal{S}$ is a critical point for φ if and only if $\boldsymbol{R}_0 \in \mathcal{D}$. In other words, the set $\mathcal{S}$ contains all the global minima and no other critical points of φ.
The proof in [11, Theorem 1] can be used provided that the following lemma is used instead of [11, Lemma 1].

Lemma 5: Let R_1, R_2, S be three rotations in $SO(3)$ such that $d(R_i, S) < \frac{\pi}{2}$, $i = 1, 2$ and let f be a reshaping function. Define the unique minimal geodesics $\gamma_i(t)$ such that $\gamma_i(0) = S$ and $\gamma_i(1) = R_i$, $i = 1, 2$. Define also $\theta_{12}(t) = d(\gamma_1(t), \gamma_2(t))$ and $\phi_{12}(t) = f(\theta_{12}(t))$. Then $\dot{\phi}_{12} \geq 0$ for $t \in (0, 1]$, with equality if and only if $R_1 = R_2$.

Proof: Notice that $\dot{\phi}_{12} = \dot{f}(\theta_{12})\dot{\theta}_{12}$. From [11, Lemma 1] we have $\dot{\theta}_{12} \geq 0$ on $(0; 1]$, and $\dot{f}(\theta_{12}) > 0$ by Def. 2. ∎

Undesired minima and θ-neighborhood sub-graphs. We give the following definition of *θ-neighborhood sub-graph*.

Definition 6: Given a graph $G = (V, E)$, a configuration $\boldsymbol{R} \in SO(3)^N$ and an angle $\theta \in [0, 2\pi)$, the θ-neighborhood sub-graph induced by $\boldsymbol{R} \in SO(3)^N$ on G is defined as $G_{\boldsymbol{R},\theta} = (V, E_{\boldsymbol{R},\theta})$, where $E_{\boldsymbol{R},\theta} \subseteq E$ is the set of edges $(i, j) \in E$ such that $\theta_{ij} = d(R_i, R_j) < \theta$.

We can relate Def. 6 to $\mathcal{S}$ and the notion of desired/undesired equilibria by defining

$$\theta_0 < \frac{\pi}{2 \operatorname{Diam}(G)}. \quad (17)$$

Proposition 7: Let $\boldsymbol{R}^\star$ be an equilibrium of (14). If $G_{\boldsymbol{R}^\star,\theta_0}$ is connected, then $\boldsymbol{R}^\star \in \mathcal{D}$, *i.e.*, it is a global minimizer. Otherwise, $\boldsymbol{R}^\star$ is an undesired equilibrium.

Proof: [Sketch] When $G_{\boldsymbol{R},\theta_0}$ is connected, one can show that all the states R_i are inside an open ball of radius $\frac{\pi}{2}$. Let $\mathcal{S}_{conn} = \{\boldsymbol{R} \in SO(3)^N : G_{\boldsymbol{R},\theta_0} \text{ is connected}\}$. Then $\mathcal{S}_{conn} \subset \mathcal{S}$. Moreover, it is easy to show that $\boldsymbol{R} \in \mathcal{D}$ implies that $G_{\boldsymbol{R},\theta_0}$ is connected, hence $\mathcal{D} \subset \mathcal{S}_{conn}$. However, the only equilibria of (14) in $\mathcal{S}$ are global minimizers, hence $\boldsymbol{R}^\star$ is an equilibrium and is in $\mathcal{S}_{conn}$ if and only if $\boldsymbol{R}^\star \in \mathcal{D}$. The rest of the claim follows easily (see also Fig. 3). ∎

In the following, we show that any equilibrium $\boldsymbol{R}$ for which $G_{\boldsymbol{R},\theta_0}$ is disconnected is unstable. Proposition 7 then will imply that the only stable equilibria are the global minimizers.

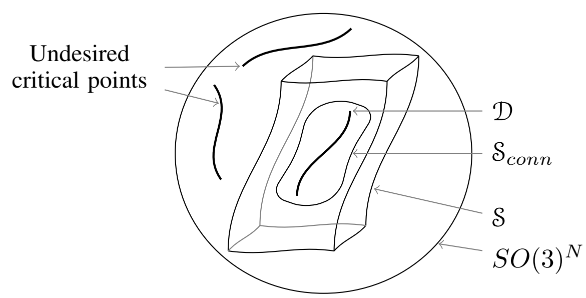


Fig. 3: Graphical description of the proof for Prop. 7

V. INSTABILITY OF UNDESIRED EQUILIBRIA

Here we show that if $\boldsymbol{R}^\star$ is an equilibrium of (14), and if $G_{\boldsymbol{R}^\star,\theta_0}$ is not connected, then $\boldsymbol{R}^\star$ cannot be stable.

We first consider non-differentiable points of φ, *i.e.*, points for which $\theta_{ij} = \pi$ for some edge $(i, j) \in E$ (see §III).

Proposition 8: Non-differentiable points of φ are not local minima.

Proof: At a given point $\boldsymbol{R}^\star \in SO(3)^N$ where φ is not differentiable, we can decompose φ as follows

$$\varphi(\boldsymbol{R}) = \sum_{(i,j)\in E_\pi} f(\theta_{ij}) + \sum_{(i,j)\in \bar{E}_\pi} f(\theta_{ij}) \doteq \varphi_\pi(\boldsymbol{R}) + \varphi_{\bar\pi}(\boldsymbol{R}), \quad (18)$$

where $E_\pi = \{(i, j) \in E : \theta_{ij}(R_i^\star, R_j^\star) = \pi\}$ and $\bar{E}_\pi = E\backslash E_\pi$. Since π is the maximum value that θ_{ij} can take, φ_π has a maximum at $\boldsymbol{R}^\star$. Next, we argue that addition of the twice differentiable term $\varphi_{\bar\pi}$ cannot turn $\boldsymbol{R}^\star$ into a local minimizer of φ. Let $\tilde{\boldsymbol{R}}^\star(\varepsilon)$ be a geodesic in $SO(3)^N$ such that $\tilde{\boldsymbol{R}}^\star(0) = \boldsymbol{R}^\star$ and $\dot{\tilde{\boldsymbol{R}}}^\star = \boldsymbol{W}$, $\boldsymbol{W} \in T_{\boldsymbol{R}^\star} SO(3)^N$. Let $\tilde\varphi(\varepsilon) = \varphi\big(\tilde{\boldsymbol{R}}^\star(\varepsilon)\big)$, $\tilde\varphi_\pi(\varepsilon) = \varphi_\pi\big(\tilde{\boldsymbol{R}}^\star(\varepsilon)\big)$, and $\tilde\varphi_{\bar\pi}(\varepsilon) = \varphi_{\bar\pi}\big(\tilde{\boldsymbol{R}}^\star(\varepsilon)\big)$. Assume that $\boldsymbol{W}$ is chosen such that $\dot{\tilde\varphi}_\pi^-(0) > \dot{\tilde\varphi}_\pi^+(0)$. Note that such $\boldsymbol{W}$ always exists, as shown by Prop. 25 in [17]. Since $\varphi_{\bar\pi}$ is differentiable at $\boldsymbol{R}^\star$, we have $\dot{\tilde\varphi}_{\bar\pi}^-(0) = \dot{\tilde\varphi}_{\bar\pi}^+(0)$, and $\dot{\tilde\varphi}^-(0) > \dot{\tilde\varphi}^+(0)$. However, a necessary condition for $\boldsymbol{R}^\star$ to be a local minimizer of φ is that $\dot{\tilde\varphi}^-(0) \le \dot{\tilde\varphi}^+(0)$ independently from the direction W. The claim follows. ■

Since non-differentiable points are not local minima, they cannot be stable.

We now consider stationary points of φ. Since the function is twice differentiable (see §III), a stationary point $\boldsymbol{R}^*$ is unstable when the Hessian of φ has at least one negative eigenvalue. The following gives a matrix form for $\operatorname{Hess}\varphi$ [17].

Proposition 9: With the identifications of $T_R SO(3)$ with $\mathbb{R}^3$ and $T_{\boldsymbol{R}} SO(3)^N$ with $\mathbb{R}^{3N}$ given by (1), the matrix representation of the Hessian of φ at $\boldsymbol{R}$ is equal to $H \in \mathbb{R}^{3N\times 3N}$ whose (k, l)-th, 3×3 block is given by

$$[H]_{k,l;3} = \begin{cases} \sum_{j:(k,j)\in E} \operatorname{sym}\big(H_{kj}\big) & \text{if } k = l \\ -H_{kl} & \text{if } (k,l) \in E, k > l \\ -H_{kl}^T & \text{if } (k,l) \in E, k < l \\ 0 & \text{otherwise} \end{cases}, \quad (19)$$

where $\operatorname{sym}(A) = \frac{1}{2}(A + A^T)$ denotes the symmetric part of a matrix, and where

$$H_{ij}(R_i, R_j) = \ddot f(\theta_{ij}) u_{ij} u_{ij}^T + \dot f(\theta_{ij}) D_{ij}, \quad (20)$$

and $D_{ij}(R_i, R_j) = \mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R_i^T R_j)$.

The following lemma characterizes the spectral decomposition of $\operatorname{sym}(H_{ij})$ (see [17] for a proof).

Lemma 10: The spectrum of $\operatorname{sym}(H_{ij})$ is given by

$$\sigma\big(\operatorname{sym}(H_{ij})\big) = \{\ddot f(\theta_{ij}), \frac{\dot f(\theta_{ij})}{\theta_{ij}}\alpha(\theta_{ij}), \frac{\dot f(\theta_{ij})}{\theta_{ij}}\alpha(\theta_{ij})\} \quad (21)$$

where $\alpha(\theta)$ is given in Prop. 1. The first eigenvector is u_{ij} and the other two eigenvectors are orthogonal to the first. Note that $\alpha(\theta) \in [0; 1]$ for $\theta \in [0; \pi]$.

The main idea for the proof of Lemma 10 is to use Prop. 1 to obtain the spectral decomposition of H_{ij}, compute its real Jordan canonical form and then use this decomposition to obtain $\sigma\big(\operatorname{sym}(H_{ij})\big)$. A detailed proof can be found in [17].

The following theorem gives sufficient conditions on the reshaping function f which imply that $\operatorname{Hess}\varphi(\boldsymbol{R})$ has a negative eigenvalue whenever $G_{\boldsymbol{R},\theta_0}$ is disconnected.

Theorem 11: Let $\boldsymbol{R} \in SO(3)^N$ be any configuration of states for which $G_{\boldsymbol{R},\theta_0}$ is not connected. Define a partition of the vertices $V = \{V_c, V_{\bar c}\}$, where V_c represents a connected component of $G_{\boldsymbol{R},\theta_0}$. Let $E_{c\bar c} = \{(i, j) \in E : i \in V_c, j \in V_{\bar c}\}$ be the set of edges between the two elements in the partition and let $q \in \mathbb{R}$ be a number satisfying

$$0 < q < \max_{v\in\mathbb{S}^2} \min_{(i,j)\in E_{c\bar c}} (v^T u_{ij})^2, \quad (22)$$

Assume that the reshaping function $f(\theta)$ satisfies the following properties:

1) $\ddot f(\theta) < 0$ for $\theta \in [\theta_0, \pi]$.
2) $\ddot f(\theta)q + \frac{\dot f(\theta)}{\theta}(1 - q) < 0$ for $\theta \in [\theta_0, \pi]$.

Then, the Hessian of φ evaluated at $\boldsymbol{R}$ has at least one negative eigenvalue.

Proof: From the Rayleigh-Ritz theorem [22], if there exist $\boldsymbol{u} \in \mathbb{R}^{3N}$ such that $\boldsymbol{u}^T H\boldsymbol{u} < 0$, then the matrix H has at least one negative eigenvalue. We now construct such $\boldsymbol{u}$.

Let $N_c = |V_c|$ and assume, without loss of generality, that $V_c = \{1, \dots, N_c\}$ (if not, just reorder the vertices). Define:

$$\boldsymbol{u} = \begin{bmatrix} u \otimes \mathbf{1}_{N_c} \\ -u \otimes \mathbf{1}_{N-N_c} \end{bmatrix} = \big[\underbrace{u^T \quad \dots \quad u^T}_{N_c \text{ times}} \underbrace{-u^T \quad \dots \quad -u^T}_{N-N_c \text{ times}}\big]^T, \quad (23)$$

where $\mathbf{1}_d$ represents the vector of all ones in $\mathbb{R}^d$, $\otimes$ represents the Kronecker's product and $u \in \mathbb{R}^3$ is defined as

$$u \doteq \arg\max_{v\in\mathbb{S}^2} \min_{(i,j)\in E_{c\bar c}} (v^T u_{ij})^2. \quad (24)$$

Note that $q < (u^T u_{ij})^2$ for all $(i, j) \in E_{c\bar c}$. The quantity $\boldsymbol{u}^T H\boldsymbol{u}$ can be computed as

$$\boldsymbol{u}^T H\boldsymbol{u} = u^T\Big(\sum_{i\in V}\sum_{j:(i,j)\in E} \operatorname{sym}(H_{ij})\Big)u \quad (25a)$$

$$+ u^T \sum_{i\in V_c}\Big(\sum_{\substack{j\in V_c\\(i,j)\in E}} -\operatorname{sym}(H_{ij}) + \sum_{\substack{j\in V_{\bar c}\\(i,j)\in E}} \operatorname{sym}(H_{ij})\Big)u \quad (25b)$$

$$+ u^T \sum_{i\in V_{\bar c}}\Big(\sum_{\substack{j\in V_c\\(i,j)\in E}} \operatorname{sym}(H_{ij}) + \sum_{\substack{j\in V_{\bar c}\\(i,j)\in E}} -\operatorname{sym}(H_{ij})\Big)u \quad (25c)$$

where (25a) corresponds to the diagonal blocks of H, (25b) comes from the first N_c row-blocks of H, (25c) comes from the remaining $(N - N_c)$ row-blocks of H and we used the fact that $u^T \operatorname{sym}(H_{ij})u = u^T H_{ij} u = u^T H_{ij}^T u$. Then, the expression in (25) can be simplified as

$$\boldsymbol{u}^T H\boldsymbol{u} = 4 \sum_{(i,j)\in E_{c\bar c}} u^T \operatorname{sym}(H_{ij})u \quad (26)$$

Note that, from the definition of $G_{\boldsymbol{R},\theta_0}$, we have $\theta_{ij} > \theta_0$ for all the edges involved in the sum (26).

From Lemma 10, we have

$$u^T H_{ij} u = \ddot{f}(\theta_{ij})(u^T u_{ij})^2 + \frac{\alpha(\theta_{ij})}{\theta_{ij}} \dot{f}(\theta_{ij})\big(1 - (u^T u_{ij})^2\big), \tag{27}$$

where the first term comes from the projection on the space spanned to the first eigenvector of H_{ij} and the second term comes from the projection on the orthogonal complement. We focus on the interval $\theta_{ij} \in [\theta_0, \pi]$. Note that, on this interval, $\alpha(\theta_{ij}) \in [0,1]$. Also, by assumption, we have $\dot{f}(\theta_{ij}) > 0$ and $\ddot{f}(\theta_{ij}) < 0$. Since $(u^T u_{ij})^2 \in [0,1]$, (27) can be seen as a convex combination of the two values $\ddot{f}(\theta_{ij})$, which is negative, and $\frac{\alpha(\theta_{ij})}{\theta_{ij}} \dot{f}(\theta_{ij})$, which is positive. From these considerations and the definition of q we have that

$$u^T H_{ij} u \leq \ddot{f}(\theta_{ij}) q + \frac{\dot{f}(\theta_{ij})}{\theta_{ij}}(1-q). \tag{28}$$

From the assumption on f we therefore conclude that $u^T H_{ij} u < 0$ for $\theta_{ij} \in [\theta_0, \pi]$, and, from (26), $\boldsymbol{u}^T H \boldsymbol{u} < 0$. The claim follows. ■

The next step in our development is to show that the proposed reshaping function (13) satisfies the assumptions of Thm. 11 when the parameter b is large enough.

VI. PROPERTIES OF THE RESHAPING FUNCTION

In this section we show that it is possible to choose a parameter b in (13) such that the only stable equilibria of the consensus protocol are desired equilibria.

We have the following result:

Proposition 12: The function (13) satisfies the assumptions of Def. 2 and Thm. 11 when $b > (q\theta_0)^{-1}$. Moreover, $f(0) = 0$, $f(\pi) = \pi^2$, $\dot{f}(0) = 0$, $\frac{\dot{f}(\theta)}{\theta} \leq ab$ and $\ddot{f}(\theta) \leq ab$ for $\theta \geq 0$.

Remark 1: The property $f(\pi) = \pi^2$ fixes a scaling for $f(\theta)$. The choice of the particular value of $f(\pi)$ is somewhat arbitrary and can be changed by scaling a accordingly.

Proof: We first show $a > 0$. The derivative of $f_0(\theta)$ is

$$\dot{f}_0(\theta) = b\theta \exp(-b\theta). \tag{29}$$

Note $f(0) = f_0(0) = \dot{f}_0(0) = 0$. By inspection, $b > 0$, and therefore $\dot{f}_0(\theta) > 0$ for $\theta > 0$. It follows that also $f(\theta) > 0$ for $\theta > 0$. Since $a = \frac{\pi^2}{2 f_0(\pi)}$, we have $a > 0$ and $f(\pi) = \pi^2$. Now, the derivatives of f are

$$\dot{f}(\theta) = a \dot{f}_0(\theta) = ab\theta \exp(-b\theta), \tag{30}$$

$$\ddot{f}(\theta) = ab(1 - b\theta)\exp(-b\theta). \tag{31}$$

From the discussion above, we have $\dot{f}(0) = 0$ and $\dot{f}(\theta) > 0$ for $\theta > 0$, which is the first property we needed to show. Also, note $b\theta_0 = q^{-1} > 1$, hence $\ddot{f}(\theta) < 0$ for $\theta > \theta_0$. This shows the second property. Next,

$$\ddot{f}(\theta) q + \frac{\dot{f}(\theta)}{\theta}(1-q) \leq ab(q - bq\theta + 1 - q)\exp(-b\theta) = ab(1 - bq\theta)\exp(-b\theta), \tag{32}$$

which is negative for $\theta > (bq)^{-1} = \theta_0$, as requested by the third property. Finally, notice that we have

$$\ddot{f}(\theta) = ab\exp(-b\theta) - ab^2\theta\exp(-b\theta) \leq ab\exp(-b\theta) \leq ab \tag{33}$$

and

$$\frac{\dot{f}(\theta)}{\theta} = ab\exp(-b\theta) \leq ab. \tag{34}$$

This concludes the proof. ■

Remark 2: One subtle issue here is that the value of b must be fixed a priori, and it can be finite only if there exist a number $q > 0$ satisfying (22) for any possible configuration $\boldsymbol{R} \in SO(3)^N$ for which $G_{\boldsymbol{R},\theta_0}$ is not connected. Luckily, one can show that, given the graph G, such number $q > 0$ always exists. The proof of this fact is rather technical and it can be found in [17]. Unfortunately, this proof is non-constructive, and we were not able to give a distributed way to compute a bound on q (and therefore on b). Hence the notion that b should be "sufficiently large" in §III, whose existence is at least ensured by our current results. We plan to investigate constructive methods for estimating q in our future research.

VII. CHOICE OF STEP SIZE AND CONVERGENCE

Tracing back all the results of §IV–VI, we can deduce that the only stable equilibria of the protocol (14) are global minimizers $\boldsymbol{R}^\star \in \mathcal{D}$, where all the states are aligned. However, this is not enough to ensure the convergence of the protocol. In this section, we give bounds on the choice of the step size ε such that the cost φ is reduced at each step. This ensures the convergence of the protocol to the set of global minimizers $\mathcal{D}$.

Our method follows and extends ideas from [11]. First, we need the following.

Proposition 13: Wherever φ is twice differentiable, the bound on the maximum eigenvalue of the Hessian $\operatorname{Hess}\varphi$ is

$$\mu_{\max} = 2\operatorname{Deg}(G) \max_{\theta\in[0,\pi]} \max\Big\{\frac{\dot{f}(\theta)}{\theta}, \ddot{f}(\theta)\Big\} = 2ab\operatorname{Deg}(G). \tag{35}$$

This proposition can be proved by starting from Lemma 10. Its result can then be used in conjunction with the following.

Theorem 14: Let $\mu_{\max}$ be a uniform bound on the Hessian of φ as given by Prop. 13. Define $\tilde{\boldsymbol{R}}_0(\varepsilon) = \exp_{\boldsymbol{R}_0} \varepsilon \operatorname{grad}_{\boldsymbol{R}} \varphi(\boldsymbol{R}_0)$ and $\tilde{\varphi}(\varepsilon) = \varphi(\tilde{\boldsymbol{R}}_0)$. Then $\tilde{\varphi}(\varepsilon) \leq \tilde{\varphi}(0)$ for $\varepsilon \in (0, 2\mu_{\max}^{-1})$, with equality if and only if $\boldsymbol{R}_0$ is a stationary point of φ.

This theorem is almost identical to [11, Thm. 3], with the exception that $\tilde{\boldsymbol{R}}$ is not restricted to a specific set. To show this result, we first need the following lemma, whose proof can be found in [17].

Lemma 15: For any $\varepsilon_0 \in \mathbb{R}$ we have $\dot{\tilde{\varphi}}^-(\varepsilon_0) \geq \dot{\tilde{\varphi}}^+(\varepsilon_0)$.

Proof: [of Theorem 14] It is possible to show (see Prop. 26 in [17]) that $\tilde{\varphi}$ can be upper bounded with the quadratic function

$$u_0(\varepsilon) = \tilde{\varphi}(0) + \dot{\tilde{\varphi}}(0)\varepsilon + \frac{-\mu_{\max}\dot{\tilde{\varphi}}(0)}{2}\varepsilon^2. \tag{36}$$

In general, this bound stops to be valid after a point where φ is not differentiable. However, in our case, we have $\dot{\tilde{\varphi}}^-(\varepsilon) \geq$

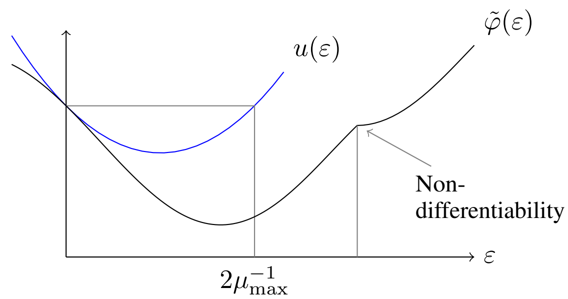


Fig. 4: The cost function φ restricted to a line and our quadratic upper bound

$\dot{\tilde{\varphi}}^{+}(\varepsilon)$ for any $\varepsilon \in \mathbb{R}$, as shown in Lemma 15. This is true in particular at all non-differentiable points of $\tilde{\varphi}$. From Prop. 26 in [17] it then follows that $u_0(\varepsilon) \geq \tilde{\varphi}(\varepsilon)$ for any $\epsilon > 0$. Again from Prop. 26 in [17], we have also that $u_0(\varepsilon) \leq u_0(0)$ for $\varepsilon \in (0, 2\mu_{\max}^{-1})$. The claim then follows. See Fig. 4 for an illustration. ■

We can finally state the main convergence theorem.

Theorem 16 (Almost-global convergence): There exist a sufficiently large $b \in \mathbb{R}$ and $\varepsilon \in \left(0; \frac{1}{ab\,\mathrm{Deg}(G)}\right)$, where a is given in §III, such that:

- The protocol (14) converges to the set of its equilibria.
- The only asymptotically stable equilibria set is set of configurations with aligned states.

Proof: [Sketch] Thm. 14 ensures that, with the given value of ε, the cost does not increase after each iteration of the protocol (14). Using a standard argument, one can show that $\boldsymbol{R}(k)$ converges to the set equilibria of (14). Then, from §IV–VI, we know that the only stable equilibria set is the set of global minimizers. ■

The presence of any numerical perturbation added to the iterates, together with Thm. 16, implies that our algorithm eventually leaves any unstable equilibrium, and converge to the set of global minimizers.

VIII. CONCLUSIONS AND FUTURE WORK

We proposed a discrete time algorithm to align the rotational states of agents connected in a network. The fundamental idea is to reshape the cost function used in Riemannian consensus [11] so that the only stable equilibria of the proposed protocol are global minimizers. We showed how *undesired* equilibria are related to θ_0*-neighborhood subgraphs*, and how the properties of the *reshaping function* $f(\theta)$ imply that the equilibria are unstable whenever such graphs are non-connected. Finally, we showed that it is possible to choose a step size such that the global cost function is reduced at every step. Combined with the instability of the undesired equilibria, this implies *almost-global convergence*, in that the only stable equilibria are global minimizers.

Unfortunately, our results depend on the choice of a parameter b which must be "sufficiently large". However, an excessively large value for b would reduce the step size and therefore the convergence speed of the algorithm. In our future research we plan to investigate a distributed method to automatically choose such parameter.

REFERENCES

[1] F. Bullo, J. Cortés, and S. Martínez, *Distributed Control of Robotic Networks*, ser. Applied Mathematics Series. Princeton University Press, 2009.
[2] A. Sarlette, S. Bonnabel, and R. Sepulchre, "Coordinated motion design on Lie groups," *IEEE Transactions on Automatic Control*, vol. to be published, 2010.
[3] T. Hatanaka, Y. Igarashi, M. Fujita, and M. Spong, "Passivity-based pose synchronization in three dimensions," *IEEE Transactions on Automatic Control*, vol. 57, no. 2, pp. 360 –375, 2012.
[4] T. Hatanaka, M. Fujita, and F. Bullo, "Vision-based cooperative estimation via multi-agent optimization," in *IEEE Conference on Decision and Control*, 2010.
[5] R. Tron, R. Vidal, and A. Terzis, "Distributed pose averaging in camera networks via consensus on SE(3)," in *International Conference on Distributed Smart Cameras*, 2008.
[6] R. Tron and R. Vidal, "Distributed image-based 3-D localization in camera sensor networks," in *Conference on Decision and Control*, 2009.
[7] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.
[8] R. Olfati-Saber, "Swarms on sphere: A programmable swarm with synchronous behaviors like oscillator networks," in *IEEE Conference on Decision and Control*, 2006, pp. 5060–5066.
[9] L. Scardovi, A. Sarlette, and R. Sepulchre, "Synchronization and balancing on the N-torus," *Systems and Control Letters*, vol. 56, no. 5, pp. 335–341, 2007.
[10] A. Sarlette and R. Sepulchre, "Consensus optimization on manifolds," *SIAM J. Control and Optimization*, vol. 48, no. 1, pp. 56–76, 2009.
[11] R. Tron, B. Afsari, and R. Vidal, "Average consensus on Riemannian manifolds with bounded curvature," in *IEEE Conference on Decision and Control*, 2011.
[12] ——, "Riemannian consensus for manifolds with bounded curvature," *IEEE Transactions on Automatic Control*, (To appear) 2012.
[13] A. Sarlette, "Geometry and symmetries in coordination control," Ph.D. dissertation, Liège University, 2009.
[14] A. Sarlette, R. Sepulchre, and N. E. Leonard, "Autonomous rigid body attitude synchronization," *Automatica*, vol. 45, no. 2, pp. 572–577, 2009.
[15] J. Thunberg, E. Montijano, and X. Hu, "Distributed attitude synchronization control," in *IEEE Conference on Decision and Control*, 2011.
[16] H. Bai, M. Arcak, and J. T. Wen, "Rigid body attitude coordination without inertial frame information," *Automatica*, vol. 44, no. 12, pp. 3170 – 3175, 2008.
[17] B. A. Roberto Tron and R. Vidal, "Results on an intrinsic consensus algorithm on SO(3) with almost-global convergence," The Johns Hopkins University, Dept. of Electrical and Computer Engineering, Tech. Rep., 2012.
[18] J. A. Bondy and U. S. R. Murty, *Graph Theory*. Springer, 2008.
[19] M. P. do Carmo, *Riemannian geometry*. Boston, MA: Birkhäuser, 1992.
[20] C. Udriste, *Convex functions and optimization methods on Riemannian manifolds*, ser. Mathematics and Applications. Kluwer Academic, 1994, vol. 297.
[21] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ: Princeton University Press, 2008.
[22] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.
[23] Y. Ma, S. Soatto, J. Košecká, and S. Sastry, *An Invitation to 3D Vision: From Images to Geometric Models*. Springer Verlag, 2003.
[24] E. Malis, "Contributions á la modélisation et á la commande en asservissement visuel," Ph.D. dissertation, Université de Rennes 1, 1998.

APPENDIX

We collect in this appendix all the results supporting the derivations and claims of the main part of the paper.

A. Exponential and logarithm map

The exponential map at the identity can be compactly computed using the well-known Rodrigues' formula, from which also the logarithm map can be deduced. More precisely, we have the following (see [23] for details).

Proposition 17: Let $\hat{v} \in \mathfrak{so}(3)$, and $R \in SO(3)$. The exponential map at the identity is given by

$$\theta = \|v\|, \tag{37}$$
$$u = \frac{v}{\theta} \tag{38}$$
$$R = \exp_I(\hat{v}) = I + \sin(\theta)\hat{u} + \big(1-\cos(\theta)\big)\hat{u}^2. \tag{39}$$

The logarithm map at the identity is given by

$$\theta = \arccos\left(\frac{\mathrm{tr}(R)-1}{2}\right), \tag{40}$$
$$u = \frac{1}{2\sin\theta}(R-R^T)^\vee \tag{41}$$
$$\hat{v} = \log_I(R) = \theta\hat{u} \tag{42}$$

The following simple proposition gives some properties of the matrix R in relation to the axis vector u

Proposition 18: We have the following properties

1) $Ru = u$.
2) $R\hat{u}^k = \hat{u}^k R$ for any $k \in \mathbb{N}$.

Proof: The first property can be proved easily by substitution using (39) and by recalling that $\hat{u}u = 0$. The second property can be proved by using again (39) and by regrouping terms. ■

B. Matrix representation for the differential of the logarithm map

In this section, we give closed form expressions for the matrix representation and the eigenvalue decomposition of $\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}$ defined in (6). Some of the calculations below are a simplification of the results in [24, Appendix A].

To start, let $u, v \in \mathbb{R}^3$, $\theta \in [0, 2\pi)$ and recall the following properties:

1) $\hat{u}^2 = uu^T - \|u\|^2 I$ and $\hat{u}^2$ is symmetric.
2) $\hat{u}\hat{v} - (\hat{u}\hat{v})^T = \hat{u}\hat{v} - \hat{v}\hat{u} = vu^T - uv^T = (\hat{u}v)^\wedge$.
3) $\hat{u}^2\hat{v} - \hat{v}\hat{u}^2 = -\big((2I+\hat{u}^2)v\big)^\wedge = -\big((I+uu^T)v\big)^\wedge$, if $\|u\| = 1$.
4) $\mathrm{tr}(\hat{u}) = 0$.
5) $\mathrm{tr}(\hat{u}\hat{v}) = 2u^T v$.
6) $\mathrm{tr}(\hat{u}^2\hat{v}) = 0$.
7) $\frac{1+\cos(\theta)}{\sin(\theta)} = \cot\left(\frac{\theta}{2}\right)$

Properties 1–5 can be easily verified by direct computation. Property 6 follows from the fact that $\hat{u}^2$ is symmetric and $\hat{v}$ is skew-symmetric. Property 7 follows from the half-angle trigonometric identity for the tangent of an angle.

We can now give a proof of Proposition 1

Proof: From the definition (6) and from the expression for $\mathrm{Log}(R)$ in (42), we have:

$$\begin{aligned}\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R)v &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{Log}(R)}{\|\mathrm{Log}(R)\|} = \dot{u} \\ &= \frac{\cos(\theta)}{2\sin^2(\theta)}\dot{\theta}(R-R^T)^\vee + \frac{1}{2\sin(\theta)}(\dot{R}-\dot{R}^T)^\vee. \end{aligned}\tag{43}$$

We now expand the necessary computations. From Rodrigues' formula (39):

$$\dot{R} = R\hat{v} = \hat{v} + \sin(\theta)\hat{u}\hat{v} + \big(1-\cos(\theta)\big)\hat{u}^2\hat{v} \tag{44}$$

$$(R-R^T)^\vee = 2\sin(\theta)u \tag{45}$$

$$\begin{aligned}(\dot{R}-\dot{R}^T)^\vee &= \big(2\hat{v} + \sin(\theta)(\hat{u}v)^\wedge + (1-\cos(\theta))(\hat{u}^2\hat{v} + \hat{v}\hat{u}^2)\big)^\vee \\ &= \big(2I + \sin(\theta)\hat{u} - (1-\cos(\theta))(2I+\hat{u}^2)\big)v \\ &= \big(2\cos(\theta)I + \sin(\theta)\hat{u} - (1-\cos(\theta))\hat{u}^2\big)v\end{aligned}\tag{46}$$

From the expression for θ in (41):

$$\begin{aligned}\dot{\theta} &= -\left(2\sqrt{1-\left(\frac{\mathrm{tr}(R)-1}{2}\right)^2}\right)^{-1}\mathrm{tr}(\dot{R}) \\ &= -\left(2\sqrt{1-\cos^2(\theta)}\right)^{-1}\mathrm{tr}(\dot{R}) \\ &= -\frac{1}{2\sin\theta}\sin(\theta)\,\mathrm{tr}(\hat{u}\hat{v}) = u^T v\end{aligned}\tag{47}$$

Substituting (45), (46) and (47) into (43), we obtain

$$\begin{aligned}\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R) &= -\frac{\cos(\theta)}{\sin(\theta)}uu^T + \frac{1}{2\sin(\theta)}(2\cos(\theta)I \\ &\quad + \sin(\theta)\hat{u} - (1-\cos(\theta))\hat{u}^2) \\ &= -\frac{\cos(\theta)}{\sin(\theta)}(uu^T - I) + \frac{1}{2}\hat{u} - \frac{1-\cos(\theta)}{2\sin(\theta)}\hat{u}^2 \\ &= \frac{1}{2}\hat{u} - \frac{\cos(\theta)}{\sin(\theta)}\hat{u}^2 - \left(\frac{1}{2\sin(\theta)} - \frac{\cos(\theta)}{2\sin(\theta)}\right)\hat{u}^2 \\ &= \frac{1}{2}\hat{u} - \frac{1+\cos(\theta)}{2\sin(\theta)}\hat{u}^2 = \frac{1}{2}\left(\hat{u} - \cot\left(\frac{\theta}{2}\right)\hat{u}^2\right) \\ &= \frac{1}{2}\left(\hat{u} - \cot\left(\frac{\theta}{2}\right)(uu^T - I)\right) \\ &= \frac{1}{2}\left(\cot\left(\frac{\theta}{2}\right)I + \hat{u} - \cot\left(\frac{\theta}{2}\right)uu^T\right),\end{aligned}\tag{48}$$

which gives the expression in the claim. Notice that we have expanded $\hat{u}^2$ in order to simplify the computations for the eigenvalue decomposition.

In this regard, let

$$U = \begin{bmatrix} u & v & \bar{v} \end{bmatrix}. \tag{49}$$

Then, the eigenvalue decomposition of $\hat{u}$ is

$$\hat{u} = U\Lambda_1 U^H \tag{50}$$

where

$$\Lambda_1 = \mathrm{diag}\left(0, \mathrm{j}, -\mathrm{j}\right), \tag{51}$$

Since uu^T is rank one, it has an eigenvalue equal to zero with double multiplicity, and its eigenvalue decomposition is not unique. However, for our purposes, we can use:

$$uu^T = U\Lambda_2 U^H, \tag{52}$$

where U is the same as before and

$$\Lambda_2 = \mathrm{diag}(1,0,0). \tag{53}$$

Also, since U is unitary, we have

$$I = UU^H. \tag{54}$$

By combining (50), (52) and (54) with (48), the eigenvalue decomposition given in the claim follow. ■

The following proposition gives a relation between $\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R)$ and the multiplication by R^T.

Proposition 19: We have the following identity:

$$R^T\,\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R) = \mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R)R^T = \mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R)^T. \tag{55}$$

Proof: The first equality can be shown from the explicit expression in Prop. 1 and by using Prop. 18 to reorder the terms. The second inequality can be show by computing the following:

$$\begin{aligned}
2\Big(\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R)^T - R^T\,\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R)\Big) &= \\
\hat{u}^T - \cot\Big(\frac{\theta}{2}\Big)(\hat{u}^2)^T - R^T\hat{u} + \cot\Big(\frac{\theta}{2}\Big)\hat{u}^2 & \\
= -\hat{u} - \cot\Big(\frac{\theta}{2}\Big)\hat{u}^2 - \big(\hat{u} - \sin(\theta)\hat{u}^2 + (1-\cos(\theta))\hat{u}^3\big) & \\
+ \cot\Big(\frac{\theta}{2}\Big)\big(\hat{u}^2 - \sin(\theta)\hat{u}^3 + (1-\cos(\theta)\big)\hat{u}^4 & \\
= -\hat{u} - \cot\Big(\frac{\theta}{2}\Big)\hat{u}^2 - \hat{u} + \sin(\theta)\hat{u}^2 + (1-\cos(\theta))\hat{u} & \\
+ \cot\Big(\frac{\theta}{2}\Big)\big(\hat{u}^2 + \sin(\theta)\hat{u} - (1-\cos(\theta))\hat{u}^2\big) & \\
= \Big(-1 - \cos(\theta) + \cot\Big(\frac{\theta}{2}\Big)\sin(\theta)\Big)\hat{u} & \\
+ \Big(\sin(\theta) - \cot\Big(\frac{\theta}{2}\Big)(1-\cos(\theta))\Big)\hat{u}^2 = 0, &
\end{aligned} \tag{56}$$

where we used the following facts

- $\hat{u}^3 = -\hat{u}$.
- $\hat{u}^4 = -\hat{u}^2$.
- $\cot\big(\frac{\theta}{2}\big)\sin(\theta) = \cos(\theta) + 1$.
- $\cot\big(\frac{\theta}{2}\big)(1-\cos(\theta)) = \frac{1-\cos^2(\theta)}{\sin(\theta)} = \sin(\theta)$.

■

C. Hessian of the cost function φ

In this section, our goal is to show Prop. 9 by computing the Hessian of φ.

Proof: As usual, we define $\boldsymbol{R}(t) = \{R_i(t)\}_{i\in V}$ as a geodesic in $SO(3)^N$ and $\tilde{\varphi}(t) = \varphi(\boldsymbol{R}(t))$ as the global cost function evaluated along this geodesic. We then obtain the matrix expression for the Hessian from the second derivative $\ddot{\varphi}(t)$. For the sake of clarity, we suppress, in the following derivations, the explicit dependence on t of θ_{ij}, R_i, $\tilde{\varphi}$ and their derivatives.

The first derivative of $\tilde{\varphi}$ is given by

$$\dot{\tilde{\varphi}} = \sum_{(i,j)\in E} \dot{f}(\theta_{ij})\dot{\theta}_{ij} \tag{57}$$

and the second derivative is given by

$$\ddot{\tilde{\varphi}} = \sum_{(i,j)\in E} \big(\ddot{f}(\theta_{ij})\dot{\theta}_{ij}^2 + \dot{f}(\theta_{ij})\ddot{\theta}_{ij}\big). \tag{58}$$

As a side note, if we made similar derivations for the Euclidean space or for the circle, we would have $\dot{\theta} = 1$ and $\ddot{\theta} = 0$.

From (9) and the chain rule, we have

$$\begin{aligned}
\dot{\theta}_{ij} &= \langle \dot{R}_i, -\frac{\log_{R_i} R_j}{\|\log_{R_i} R_j\|}\rangle + \langle \dot{R}_j, -\frac{\log_{R_i} R_j}{\|\log_{R_j} R_i\|}\rangle \\
&= -v_i^T\,\mathrm{Log}(R_i^T R_j) - v_j^T(R_j^T R_i) \\
&= (v_j - v_i)^T\,\mathrm{Log}(R_i^T R_j) = (v_j - v_i)u_{ij},
\end{aligned} \tag{59}$$

where $v_i = (R_i)^\wedge$ and where we used the property $\mathrm{Log}(R) = -\mathrm{Log}(R^T)$.

Before passing to the computation of $\ddot{\theta}_{ij}$, we need an expression for the tangent of the curve $R_i^T R_j$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}(R_i^T R_j) &= \dot{R}_i^T R_j + R_i^T \dot{R}_j = \hat{v}_i^T R_i^T R_j + R_i^T R_j \hat{v}_j \\
&= -R_i^T R_j (R_i^T R_i v_i)^\wedge + R_i^T R_j \hat{v}_j
\end{aligned} \tag{60}$$

where we used the fact that $R^T\hat{v}R = (R^T v)^\wedge$. In vector coordinate representation, we then have

$$\Big(\frac{\mathrm{d}}{\mathrm{d}t}(R_i^T R_j)\Big)^\vee = v_j - R_j^T R_i v_i. \tag{61}$$

Then, we have

$$\begin{aligned}
\ddot{\theta}_{ij} &= (v_j - v_i)^T \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{Log}(R_i^T R_j)}{\|\mathrm{Log}(R_i^T R_j)\|} \\
&= (v_j - v_i)^T\,\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R_i^T R_j)(v_j - R_j^T R_i v_i) \\
&= (v_j - v_i)^T D_{ij}(v_j - R_j^T R_i v_i) \\
&= \begin{bmatrix} v_i \\ v_j \end{bmatrix}^T \begin{bmatrix} D_{ij}R_j^T R_i & -D_{ij}R_j^T R_i \\ -\big(D_{ij}\big)^T & D_{ij} \end{bmatrix} \begin{bmatrix} v_i \\ v_j \end{bmatrix} \\
&= \begin{bmatrix} v_i \\ v_j \end{bmatrix}^T \begin{bmatrix} \mathrm{sym}(D_{ij}) & -D_{ij} \\ -D_{ij}^T & \mathrm{sym}(D_{ij}) \end{bmatrix} \begin{bmatrix} v_i \\ v_j \end{bmatrix},
\end{aligned} \tag{62}$$

where we used Prop. 19.

Plugging in (59) and (62) into (58), we get

$$\begin{aligned}
\ddot{\tilde{\varphi}} = \sum_{(i,j)\in E} \begin{bmatrix} v_i \\ v_j \end{bmatrix}^T \Bigg(\ddot{f}(\theta_{ij}) \begin{bmatrix} u_{ij}u_{ij}^T & -u_{ij}u_{ij}^T \\ -u_{ij}u_{ij}^T & u_{ij}u_{ij}^T \end{bmatrix} \\
+ \dot{f}(\theta_{ij}) \begin{bmatrix} \mathrm{sym}(D_{ij}) & -D_{ij} \\ -D_{ij}^T & \mathrm{sym}(D_{ij}) \end{bmatrix}\Bigg) \begin{bmatrix} v_i \\ v_j \end{bmatrix},
\end{aligned} \tag{63}$$

from which the claim follows. ■

D. Proof of Lemma 10

In this section we give a proof of Lemma 10, which details the spectral decomposition of $\mathrm{sym}(H_{ij})$. Before giving such proof, however, we need the following lemma.

Lemma 20: Then the Real Jordan Canonical Form (RJCF) of $\tilde{H}_{12}$ can be obtained as $\tilde{H}_{12} = SJS^T$ where $S, J \in \mathbb{R}^{3\times 3}$ and

$$S = \begin{bmatrix} u_{12} & \sqrt{2}\Re(v_{12}) & \sqrt{2}\Im(v_{12}) \end{bmatrix} \tag{64}$$

$$J = \begin{bmatrix} \ddot{f}(\theta_{12}) & 0 & 0 \\ 0 & \alpha(\theta_{12})\frac{\dot{f}(\theta_{12})}{\theta_{12}} & \beta(\theta_{12})\frac{\dot{f}(\theta_{12})}{\theta_{12}} \\ 0 & -\beta(\theta_{12})\frac{\dot{f}(\theta_{12})}{\theta_{12}} & \alpha(\theta_{12})\frac{\dot{f}(\theta_{12})}{\theta_{12}} \end{bmatrix} \tag{65}$$

Proof: Let $\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R_i^T R_j) = U\Lambda U^H$ be the eigenvalue decomposition of $\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R_i^T R_j)$, where

$$U = \begin{bmatrix} u_{ij} & v_{ij} & \bar{v}_{ij} \end{bmatrix} \tag{66}$$

and

$$\Lambda = \mathrm{diag}\big(0, \alpha(\theta_{ij}) + \mathrm{j}\beta(\theta_{ij}), \alpha(\theta_{ij}) - \mathrm{j}\beta(\theta_{ij})\big), \tag{67}$$

see Prop. 1. Combining the eigenvalue decomposition of $\mathrm{D}\frac{\mathrm{Log}}{\|\mathrm{Log}\|}(R_i^T R_j)$ with the definition of H_{ij} in (20), we have the eigenvalue decomposition

$$H_{ij} = U\,\mathrm{diag}\big(\sigma(H_{ij})\big)U^H, \tag{68}$$

where

$$\sigma(H_{ij}) = \{\ddot{f}(\theta_{ij}), \frac{\dot{f}(\theta_{ij})}{\theta_{ij}}\big(\alpha(\theta_{ij}) \pm \mathrm{j}\beta(\theta_{ij})\big)\} \tag{69}$$

Define

$$P = \frac{1}{\sqrt{2}}\begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & \mathrm{j} \\ 0 & 1 & -\mathrm{j} \end{bmatrix} \in \mathbb{C}^{3\times 3}. \tag{70}$$

Notice that P is unitary, *i.e.*, $PP^H = P^H P = I$. Also, $S = UP$ is real and $S^H = S^T$. Then, we can write $\tilde{H}_{12} = UPP^H \Lambda_{H_{ij}} PP^H U^H$, from which the given RJCF follows with $S = UP$ and $J = P^H \Lambda_{H_{ij}} P$. ∎

Notice that this decomposition depends only on the angle θ_{ij} and axis u_{ij} of $R_i^T R_j$, and not on the individual values of the rotations.

We can now give the proof of Lemma 10. *Proof:* Using the RJCF given in Lemma 20, the eigenvalue decomposition of $\mathrm{sym}(\tilde{H}_{12})$ can be computed as

$$\begin{aligned}\mathrm{sym}(\tilde{H}_{12}) &= \frac{1}{2}(SJS^T + SJ^T S^T) = S\,\mathrm{sym}(J)S^T \\ &= S\,\mathrm{diag}\Big(\ddot{f}(\theta_{ij}), \frac{\dot{f}(\theta_{ij})}{\theta_{ij}}\alpha(\theta_{ij}), \frac{\dot{f}(\theta_{ij})}{\theta_{ij}}\alpha(\theta_{ij})\Big)S^T. \end{aligned} \tag{71}$$

The claim follows. ∎

E. Proof of Prop. 13

Proof: Using the result of Prop. 9, the Hessian of the cost function between only two nodes $\varphi_{ij} = f(\theta_{ij})$ is given by

$$\begin{bmatrix} \mathrm{sym}(H_{ij}) & -H_{ij}^T \\ -H_{ij} & \mathrm{sym}(H_{ij}) \end{bmatrix}. \tag{72}$$

Using the RJCF decomposition of H_{ij} from the proof of Lemma 10, we have the following similarity transformations:

$$\begin{aligned}
&\begin{bmatrix} \mathrm{sym}(H_{ij}) & -H_{ij}^T \\ -H_{ij} & \mathrm{sym}(H_{ij}) \end{bmatrix} \sim \begin{bmatrix} S\,\mathrm{sym}(J)S^T & -SJ^T S^T \\ -SJ^T S^T & S\,\mathrm{sym}(J)S^T \end{bmatrix} \\
&\sim \tilde{S}\begin{bmatrix} \ddot{f} & & -\ddot{f} & \\ & \frac{\dot{f}}{\theta_{ij}}\begin{bmatrix}\alpha & \\ & \alpha\end{bmatrix} & & \frac{\dot{f}}{\theta_{ij}}\begin{bmatrix}-\alpha & \beta \\ -\beta & \alpha\end{bmatrix} \\ -\ddot{f} & & \ddot{f} & \\ & \frac{\dot{f}}{\theta_{ij}}\begin{bmatrix}-\alpha & -\beta \\ \beta & \alpha\end{bmatrix} & & \frac{\dot{f}}{\theta_{ij}}\begin{bmatrix}\alpha & \\ & \alpha\end{bmatrix} \end{bmatrix}\tilde{S} \\
&\sim \begin{bmatrix} \ddot{f}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} & \\ & \frac{\dot{f}}{\theta_{ij}}\begin{bmatrix} \alpha & & -\alpha & \beta \\ & \alpha & -\beta & \alpha \\ -\alpha & -\beta & \alpha & \\ \beta & -\alpha & & \alpha \end{bmatrix} \end{bmatrix} \\
&\doteq \begin{bmatrix} \ddot{f}K_1 & \\ & \frac{\dot{f}}{\theta_{ij}}K_2 \end{bmatrix},
\end{aligned} \tag{73}$$

where $\tilde{S} = \begin{bmatrix} S & \\ & S \end{bmatrix}$ and the dependence of $\dot{f}$, $\ddot{f}$, α and β from θ_{ij} has been omitted for brevity. The eigenvalues of the Hessian are given by the union of the eigenvalues of K_1 times $\ddot{f}$ and the eigenvalues of K_2 times $\frac{\dot{f}}{\theta_{ij}}$. The spectrum of K_1 is $\sigma(K_1) = \{2, 0\}$. The spectrum of K_2 (which can be computed with a symbolic math software) is $\sigma(K_2) = \{\alpha \pm \sqrt{\alpha^2 + \beta^2}\}$, where each eigenvalue has double multiplicity. One can then directly verify that all the eigenvalues of K_1 and K_2 are bounded by 2. This means that the eigenvalues of (72) can be bounded by $2\max_{\theta\in[0,\pi]} \max\{\frac{\dot{f}(\theta)}{\theta}, \ddot{f}(\theta)\}$. Using [12, Thm. 8] and Prop. 12, the claim follows. ∎

F. Existence of q satisfying (22)

In this section we show that it is always possible to choose a constant $q \in \mathbb{R}$ satisfying (22). We first need the following preliminary lemma.

Lemma 21: Let $\{u_1, \ldots, u_{N_u}\}$ be a finite set of N_u vectors in $\mathbb{S}^2$. There always exists a vector $u \in \mathbb{S}^2$ such that $(u_i^T u)^2 > 0$ for all $i \in \{1, \ldots, N_u\}$.

Proof: First, we denote as $\{\mathcal{S}_i\}_{i=1}^{N_u}$ the set of two-dimensional hyperplanes with normals $\{u_i\}_{i=1}^{N_u}$. Since the quantity $(u_i^T u)^2$ is always non-negative, the only case we need to examine is when $u_i^T u = 0$ for some i. This would imply that $u \in \mathcal{S}_i$. However, for finite N_u, $\bigcup_{i=1}^{N_u} \mathcal{S}_i \subset \mathbb{R}^3$, *i.e.*, the union of the hyperplanes cannot cover the entire space. Therefore, it is always possible to find a $u \notin \bigcup_{i=1}^{N_u} \mathcal{S}_i$, *i.e.*, $u \notin \mathcal{S}_i\ \forall i \in \{1, \ldots, N_u\}$, and the lemma is proven. ∎

This lemma ensures that the quantity

$$\max_{v\in\mathbb{S}^2} \min_{(i,j)\in E_{c\bar{c}}} (v^T u_{ij})^2 \tag{74}$$

is always strictly positive.

The following theorem justifies the claims made in Remark 2.

Theorem 22: Let $G = (V, E)$ be a bipartite graph where $V = (V_1, V_2)$ and E contains edges (i, j) such that $i \in V_1$ and $j \in V_2$. Imagine that for each vertex $i \in V$ we can associate a rotation R_i. Given an angle $\theta_0 \in (0, \pi]$, define the set $\mathcal{S}_{\theta_0} \subseteq SO(3)^{|V|}$ as

$$\mathcal{S}_{\theta_0}^{|V|} = \{\{R_i\}_{i \in V} \in SO(3)^{|V|} : d(R_i, R_j) \geq \theta_0 \ \forall (i,j) \in E\} \quad (75)$$

and, similarly, the set

$$\mathcal{S}_{\theta_0}^{2} = \{\{R_1, R_2\} \in SO(3)^2 : d(R_i, R_j) \geq \theta_0\}. \quad (76)$$

For each $(i, j) \in E$, define the distance $\theta_{ij}(R_i, R_j) = d(R_i, R_j) \in (\theta_0, \pi]$ and the unit-norm vector

$$u_{ij}(R_i, R_j) = \frac{\operatorname{Log}(R_i^T R_j)}{\theta_{ij}(R_i, R_j)} = \frac{1}{\sin(\theta_{ij})} \operatorname{asym}(R_i^T R_j)^\vee, \quad (77)$$

when $\theta_{ij} \in [\theta_0, \pi)$. Note that u_{ij} is not defined when $\theta_{ij} = \pi$. Define also the functions

$$g_{ij}(u, R_i, R_j) = (u^T u_{ij}(R_i, R_j))^2. \quad (78)$$

Again, note that $g_{ij}(u, R_i, R_j)$ is not defined when $d(R_i, R_j) = \pi$.

Then, we have the following:

1) We can define functions $\{\bar{g}_{ij}(u, R_i, R_j)\}_{(i,j) \in E}$ where each one is defined and continuous on the entire set $\mathbb{S}^2 \times \mathcal{S}_{\theta_0}^2$ and it is the extension by continuity of $\{g_{ij}(u, R_i, R_j)\}_{(i,j) \in E}$.
2) There exist a strictly positive constant $q \in (0, 1]$ such that

$$0 < q \leq \min_{\boldsymbol{R} \in \mathcal{S}_{\theta_0}^{|V|}} \max_{u \in \mathbb{S}^2} \min_{(i,j) \in E} \bar{g}_{ij}(u, R_i, R_j). \quad (79)$$

Proof: First, notice that $g_{ij}(u, R_i, R_j) = g_{ij}(u, I, R_{ij})$, where $R_{ij} = R_i^T R_j$. Then, g_{ij} can be written as an explicit function of R_{ij} and u as follows.

$$\begin{aligned} g_{ij}(u, I, R_{ij}) &= \left(\frac{u^T \operatorname{asym}(R_{ij})^\vee}{\sin(\theta_{ij}(R_i j))}\right)^2 \\ &= \left(\frac{-\operatorname{tr}(\hat{u} R_{ij})}{2 \sin\left(\arccos\left(\frac{\operatorname{tr}(R_{ij}) - 1}{2}\right)\right)}\right)^2 \\ &= \frac{\operatorname{tr}^2(\hat{u} R_{ij})}{4\left(1 - \left(\frac{\operatorname{tr}(R_{ij}) - 1}{2}\right)^2\right)}, \end{aligned} \quad (80)$$

where we used the fact that

$$\theta_{ij} = \arccos\left(\frac{\operatorname{tr}(R_{ij}) - 1}{2}\right) \quad (81)$$

and $\sin(\arccos(x)) = \sqrt{1 - x^2}$.

The only problems with the continuity of g_{ij} on $\mathcal{S}_{\theta_0}^2$ appear when $R_{ij} = R_0$ where $d(I, R_0) = \pi$, for which the function does not exist. We want to show that these points are in fact removable singularities.

In order to do so, let $R(t)$ be a curve in $SO(3)$ with $R(0) = R_0$ and $\dot{R}(0) = R_0 \hat{v}$, where $v \in \mathbb{R}^3$. Our strategy is to show that $\lim_{t \to 0} g_{ij}(u, I, R(t))$ does not depend on the particular path $R(t)$. By definition, this implies that $\bar{g}_{ij}(u, I, R_{ij})$ is continuous.

We will denote as $u_0 \in \mathbb{S}^2$ a vector such that $R_0 = \exp_I(\pi \hat{u}_0)$. Note that we could also choose $-u_0$. However, this choice does not influence the final result, as we will see later. From Rodrigues' formula, we have that $R_0 = -I + 2u_0 u_0^T$ and, hence, R_0 is symmetric. In order to obtain an expression for the Taylor for the numerator of (80), consider first the following function:

$$g_{N0}(t) = \operatorname{tr}(\hat{u} R(t)). \quad (82)$$

Note that $g_{N0}(0) = \operatorname{tr}(\hat{u} R_0) = 0$, because $\hat{u}$ is skew-symmetric and R_0 is symmetric. The derivative of $g_{N0}(t)$ is

$$\dot{g}_{N0}(t) = \operatorname{tr}(\hat{u} R(t) \hat{v}) \quad (83)$$

Now, the numerator of (80) evaluated along $R(t)$ is given by

$$g_N(t) = g_{N0}^2(t) = \overset{2}{\operatorname{tr}}(\hat{u} R(t)) \quad (84)$$

The first two derivatives of $g_N(t)$ are given by

$$\dot{g}_N(t) = 2 g_{N0}(t) \dot{g}_{N0}(t), \quad (85)$$

$$\ddot{g}_N(t) = 2\left(\dot{f}_{N0}^2(t) + g_{N0}(t) \ddot{g}_{N0}(t)\right). \quad (86)$$

For $t = 0$ these reduce to

$$g_N(0) = \dot{g}_N(0) = 0, \quad (87)$$

$$\begin{aligned} \ddot{g}_N(0) &= \overset{2}{\operatorname{tr}}(\hat{u} R_0 \hat{v}) = \overset{2}{\operatorname{tr}}(\hat{u}(-I + 2u_0 u_0^T) \hat{v}) \\ &= \left(-\operatorname{tr}(\hat{u}\hat{v}) + 2 \operatorname{tr}(\hat{u} u_0 u_0^T \hat{v})\right)^2 \\ &= \left(2u^T v + 2u_0^T \hat{u} \hat{v} u_0)\right)^2 = 4\left(u^T v + u^T \hat{u}_0^2 v)\right)^2 \\ &= 4\left(u^T v + u^T(-I + u_0 u_0^T) v)\right)^2 \\ &= 4(u^T u_0)^2 (v^T u_0)^2. \end{aligned} \quad (88)$$

Now we consider the denominator of (80). Similarly to before, define

$$g_{D0}(t) = \frac{\operatorname{tr}(R(t)) - 1}{2}. \quad (89)$$

The first two derivatives of $g_{D0}(t)$ are given by

$$\dot{g}_{D0}(t) = \frac{1}{2} \operatorname{tr}(\hat{v} R(t)), \quad (90)$$

$$\ddot{g}_{D0}(t) = \frac{1}{2} \operatorname{tr}(\hat{v}^2 R(t)). \quad (91)$$

When $t = 0$ these reduce to

$$g_{D0}(0) = -1, \quad (92)$$

$$\dot{g}_{D0}(0) = 0, \quad (93)$$

$$\begin{aligned} \ddot{g}_{D0}(0) &= \frac{1}{2} \operatorname{tr}\left(\hat{v}^2(-I + 2u_0 u_0^T)\right) \\ &= -\frac{1}{2} \operatorname{tr} \hat{v}^2 + u_0^T(-\|v\|^2 I + vv^T) u_0 \\ &= \|v\|^2 - \|v\|^2 \|u_0\|^2 + (v^T u_0)^2 \\ &= (v^T u_0)^2. \end{aligned} \quad (94)$$

The denominator of (80) evaluated along $R(t)$ and its derivatives are given by

$$g_D(t) = 1 - g_{D0}^2(t), \quad (95)$$
$$\dot{g}_D(t) = 2g_{N0}(t)\dot{g}_{N0}(t), \quad (96)$$
$$\ddot{g}_D(t) = 2\big(\dot{f}_{N0}^2(t) + g_{N0}(t)\ddot{g}_{N0}(t)\big). \quad (97)$$

For $t = 0$ these reduce to

$$g_D(0) = \dot{g}_D(0) = 0, \quad (98)$$
$$\ddot{g}_D(0) = 2(v^T u_0)^2. \quad (99)$$

We can use (87) and (98) to express the ratio in (80) computed along $R(t)$ using Taylor series, and compute the limit

$$\lim_{t\to 0} g_{ij}\big(u, I, R(t)\big) = \frac{4(u^T u_0)^2 (v^T u_0)^2 t^2 + \mathrm{o}(t^2)}{2(v^T u_0)^2 t^2 + \mathrm{o}(t^2)} = 2(u^T u_0)^2 \doteq \bar{g}_{ij}(u, I, R_0). \quad (100)$$

As anticipated, this value does not depend on the particular path followed by $R(t)$, and is also independent on the choice of sign for u_{ij}. Therefore, $\bar{g}_{ij}$ is continuos at R_0. The first claim of the lemma follows.

Regarding the second claim, from Lemma 21 we know that

$$\max_{u\in\mathbb{S}^2} \min_{(i,j)\in E} \bar{g}_{ij}(u, R_i, R_j) > 0, \quad (101)$$

which is equivalent to (74). Thanks to the first part of the proof, we also know that this quantity is continuous as a function of R_i, R_j. Combined with the fact that $\mathbb{S}_{\theta_0}^{|V|}$ is compact, we deduce that the minimizer of the RHS of (79) is in the set, and the corresponding cost is positive. Hence, we can always pick q' such that (79) is satisfied. ■

G. Maxima of the reshaped distance function along geodesics

Consider the function

$$\tilde{f}_{ij}(t) = f\Big(\theta_{ij}\big(R_i(t), R_j(t)\big)\Big) = f\Big(d\big(I, R_{ij}(t)\big)\Big) \quad (102)$$

where f is a reshaping function, $R_{ij}(t) = R_i(t)^T R_j(t)$, $R_i = R_{i0}\exp(t\hat{v}_i)$ and $R_j = R_{j0}\exp(t\hat{v}_j)$ are two geodesics in $SO(3)$ starting from rotations R_{i0} and R_{j0} and with tangents v_i and v_j (using the identification of (1)). Note that the tangent of R_{ij} is given by

$$\dot{R}_{ij}(t) = \dot{R}_i(t)^T R_j(t) + R_i(t)^T \dot{R}_j(t) = \hat{v}_i^T R_{ij}(t) + R_{ij}(t)\hat{v}_j = R_{ij}(t)\big(-R_{ij}(t)^T\hat{v}_i R_{ij}(t) + \hat{v}_j\big) = R_{ij}(t)\hat{v}_{ij}(t), \quad (103)$$

where $v_{ij}(t) = v_j - R_{ij}(t)^T v_i$. In the following, we use the shorthand notation $v_{ij0} = v_{ij}(0)$. The function $\tilde{f}_{ij}$ is differentiable everywhere, except at the points where $d(I, R_{ij}(t)) = \pi$, for which it is continuous but not differentiable. In this section, we want to show that at these point the function has, in general, a strict maximum. More in detail, assume that $R_{ij}(0) = R_{i0}^T R_{j0} \doteq R_{ij0}$ satisfies $d(I, R_{ij0}) = 0$. Then, we have the following proposition.

Proposition 23: With the notation defined above, let $u_{ij0} \in \mathbb{R}^3$ be a vector satisfying $R_{ij0} = \exp(\pi\hat{u}_{ij0})$. Then

$$\dot{\tilde{f}}_{ij}^-(0) = \sqrt{2}|v_{ij0}^T u_{ij0}|\dot{f}(\pi) \quad (104)$$
$$\dot{\tilde{f}}_{ij}^+(0) = -\sqrt{2}|v_{ij0}^T u_{ij0}|\dot{f}(\pi) \quad (105)$$
$$(106)$$

Note that $\dot{\tilde{f}}_{ij}^-(0) \geq 0 \geq \dot{\tilde{f}}_{ij}^+(0)$ with equality if and only if $v_{ij0}^T u_{ij0} = 0$.

Proof: We start from the closed formula for θ_{ij} given in (81) and we compute the left and right derivatives $\dot{\theta}_{ij}^-(0)$ and $\dot{\theta}_{ij}^+(0)$. Following the definition and computations in (89)–(94), the Taylor expansion of the argument of $\arccos$ in (98) around $t = 0$ is given by:

$$g_{D0}(t) = -1 + (v_{ij0}^T u_{ij0})^2 t^2 + \mathrm{o}(t^2). \quad (107)$$

The Taylor expansion of $\arccos$ around $t = -1$ is given by

$$\arccos(t) = \pi - \sqrt{2}\sqrt{t-1} + \mathrm{o}(\sqrt{t-1}). \quad (108)$$

We can combine the two Taylor expansions in the following limits

$$\lim_{t\to 0^\pm} \frac{\theta_{ij}(t) - \theta_{ij}(0)}{t} = \lim_{t\to 0^\pm} \frac{-\sqrt{2}\sqrt{(v_{ij0}^T u_{ij0})^2 t^2} + \mathrm{o}(|t|)}{t} = \lim_{t\to 0^\pm} \frac{-\sqrt{2}|v_{ij0}^T u_{ij0}||t| + \mathrm{o}(|t|)}{t}. \quad (109)$$

The left and right derivatives of θ_{ij} around $t = 0$ are therefore

$$\dot{\theta}_{ij}^-(0) = \lim_{t\to 0^-} \frac{\theta_{ij}(t) - \theta_{ij}(0)}{t} = \sqrt{2}|v_{ij0}^T u_{ij0}| \quad (110)$$
$$\dot{\theta}_{ij}^+(0) = \lim_{t\to 0^+} \frac{\theta_{ij}(t) - \theta_{ij}(0)}{t} = -\sqrt{2}|v_{ij0}^T u_{ij0}| \quad (111)$$

The claim then easily follows. ■

The result of this lemma can be extended with the following corollary.

Corollary 24: We can always choose v_i and v_j so that $\dot{\tilde{f}}_{ij}^-(0) > 0 > \dot{\tilde{f}}_{ij}^+(0)$ with a strict inequality.

Proof: In the proof of Prop. 23, pick $v_i = 0$ and $v_j = u_{ij0}$. The claim then follows. ■

We now show that this corollary is in fact true for the more general case with $N > 2$ rotations. Expanding the notation defined in §II and in the proof of Prop. 8, let $\tilde{\varphi}_\pi(t) = \varphi_\pi\big(\tilde{\boldsymbol{R}}^\star(\varepsilon)\big)$. Then, we have the following proposition

Proposition 25: There always exists a tangent vector $\boldsymbol{W} \in T_{\boldsymbol{R}^\star}SO(3)^N$ such that $\dot{\tilde{\varphi}}_\pi^-(0) > \dot{\tilde{\varphi}}_\pi^+(0)$ with strict inequality. This implies that $\tilde{\varphi}(0) > \tilde{\varphi}(t)$ for $t \neq 0$ and t in a neighborhood of zero.

Proof: Let $\boldsymbol{W} = \{w_i\}_{i\in V}$. For an arbitrary $(\hat{\imath}, \hat{\jmath}) \in E_\pi$, pick $w_{\hat{\imath}}$ and $w_{\hat{\jmath}}$ so that $\dot{\tilde{f}}_{\hat{\imath}\hat{\jmath}}^-(0) > \dot{\tilde{f}}_{\hat{\imath}\hat{\jmath}}^+(0)$ (see Corol. 24). Since $\dot{\tilde{\varphi}}_\pi^\pm(0) = \sum_{(i,j)\in E_\pi} \dot{\tilde{f}}_{ij}^\pm(0)$, and from Prop. 23, the claim follows. ■

H. Proof of Lemma 15

Proof: If $\tilde{\varphi}$ is differentiable at ε_0, then $\dot{\tilde{\varphi}}^-(\varepsilon_0) = \dot{\tilde{\varphi}}^+(\varepsilon_0)$ by definition. Then, assume that $\tilde{\varphi}$ is not differentiable at ε_0, and define $\tilde{\varphi}_\pi(\varepsilon) = \varphi_\pi(\tilde{\boldsymbol{R}}_0)$ and $\tilde{\varphi}_{\bar{\pi}}(\varepsilon) = \varphi_{\bar{\pi}}(\tilde{\boldsymbol{R}}_0)$, where φ_π and $\varphi_{\bar{\pi}}$ are defined as in (18). From Prop. 23 we have $\dot{\tilde{\varphi}}_\pi^-(\varepsilon_0) \geq \dot{\tilde{\varphi}}_\pi^+(\varepsilon_0)$. At the same time, since $\tilde{\varphi}_{\bar{\pi}}$ is differentiable at ε_0, we have $\dot{\tilde{\varphi}}_{\bar{\pi}}^-(\varepsilon_0) = \dot{\tilde{\varphi}}_{\bar{\pi}}^+(\varepsilon_0)$. The claim then follows from the fact that $\tilde{\varphi}(\varepsilon) = \tilde{\varphi}_\pi(\varepsilon) + \tilde{\varphi}_{\bar{\pi}}(\varepsilon)$. ■

I. Quadratic upper bounds

Proposition 26: Let $\tilde{\varphi}(\varepsilon)$ be a continuous function which is twice differentiable everywhere on $\mathbb{R}^+$ except for a countable ordered set of isolated points $\{\varepsilon_2, \varepsilon_3, \ldots\} \subset \mathbb{R}$, $0 < \varepsilon_2 < \varepsilon_3$, etc. By convention, let $\varepsilon_1 = 0$. Assume $\ddot{\tilde{\varphi}}(\varepsilon) \leq \tilde{\mu}_{\max}$ for all $\varepsilon > 0$ where $\tilde{\varphi}(\varepsilon)$ is twice differentiable and with $\tilde{\mu}_{\max} > 0$. Define

$$u_{\varepsilon_0}(\varepsilon) \doteq \tilde{\varphi}(\varepsilon_0) + \dot{\tilde{\varphi}}^+(\varepsilon_0)(\varepsilon - \varepsilon_0) + \frac{\tilde{\mu}_{\max}}{2}(\varepsilon - \varepsilon_0)^2, \quad (112)$$

with $\varepsilon_0 > 0$ arbitrary.

1) We have the bounds $\tilde{\varphi}(\varepsilon) \leq u_{\varepsilon_0}(\varepsilon)$ and $\dot{\tilde{\varphi}}(\varepsilon) \leq \dot{u}_{\varepsilon_0}(\varepsilon)$ for $\varepsilon \in [\varepsilon_0; \min_{i:\varepsilon_i > \varepsilon_0} \varepsilon_i]$.
2) If $\dot{\tilde{\varphi}}^+(\varepsilon_0) < 0$, then $u_{\varepsilon_0}(\varepsilon) \leq u_{\varepsilon_0}(\varepsilon_0)$ for $\varepsilon \in [\varepsilon_0; \varepsilon_0 - \frac{2\dot{\tilde{\varphi}}^+(\varepsilon_0)}{\tilde{\mu}_{\max}}]$.
3) $u_{\varepsilon_0}(\varepsilon) = u_{\varepsilon_0}(\varepsilon^\star) + \dot{u}_{\varepsilon_0}(\varepsilon^\star)(\varepsilon - \varepsilon^\star) + \frac{\tilde{\mu}_{\max}}{2}(\varepsilon - \varepsilon^\star)^2$ for any $\varepsilon, \varepsilon^\star \in \mathbb{R}$.
4) If $\dot{\tilde{\varphi}}^+(\varepsilon_i) \leq \dot{\tilde{\varphi}}^-(\varepsilon_i)$ with $i \in \{2, 3, \ldots\}$, then $\tilde{\varphi}(\varepsilon) \leq u_{\varepsilon_{i-1}}(\varepsilon)$ for $\varepsilon \in [\varepsilon_{i-1}; \varepsilon_{i+1}]$.

Proof: We prove each claim separately.

1) By direct computation we have $u_{\varepsilon_0}(\varepsilon_0) = \tilde{\varphi}(\varepsilon_0)$, $\dot{u}_{\varepsilon_0}(\varepsilon_0) = \dot{\tilde{\varphi}}(\varepsilon_0)$ and $\ddot{u}_{\varepsilon_0}(\varepsilon) \geq \ddot{\tilde{\varphi}}(\varepsilon)$, for all $\varepsilon \in [\varepsilon_0; \min_{i:\varepsilon_i > \varepsilon_0} \varepsilon_i]$. Define $r(\varepsilon) = u_{\varepsilon_0}(\varepsilon) - \tilde{\varphi}(\varepsilon)$. Then r is convex because $\ddot{r}(\varepsilon) \geq 0$. Also, $\dot{r}(\varepsilon_0) = 0$, hence r reaches its minimum $r(\varepsilon_0) = 0$ for $\varepsilon = \varepsilon_0$. This implies $r(\varepsilon) \geq 0$ and $\dot{r}(\varepsilon) \geq 0$ for all $\varepsilon \in [\varepsilon_0; \min_{i:\varepsilon_i > \varepsilon_0} \varepsilon_i]$, and the claim follows.
2) Let ε_{min} be the minimizer of $u_{\varepsilon_0}(\varepsilon)$. Since u is a convex parabola and $\dot{u}_{\varepsilon_0}(\varepsilon_0) = \dot{\tilde{\varphi}}^+(\varepsilon_0) < 0$, we have $\varepsilon_{min} > 0$. Also, since parabolas are symmetric around the minimizer, we can deduce that $\tilde{\varphi}(\varepsilon) < \tilde{\varphi}(\varepsilon_0)$ for $\varepsilon \in \big(\varepsilon_0, \varepsilon_0 + 2(\varepsilon_{min} - \varepsilon_0)\big)$ (see Figure 4). From the condition $\dot{u}_{\varepsilon_0}(\varepsilon_{min}) = 0$ we get $\varepsilon_{min} = \varepsilon_0 - \frac{\dot{\tilde{\varphi}}^+(\varepsilon_0)}{\tilde{\mu}_{\max}}$
3) Can be verified by substitution and direct computation.
4) Notice that

$$\begin{aligned} &u_{\varepsilon_{i-1}}(\varepsilon) = \\ &u_{\varepsilon_{i-1}}(\varepsilon_i) + \dot{u}_{\varepsilon_{i-1}}(\varepsilon_i)(\varepsilon - \varepsilon_i) + \frac{\tilde{\mu}_{\max}}{2}(\varepsilon - \varepsilon_i)^2 \\ &\geq \varphi(\varepsilon_i) + \dot{\tilde{\varphi}}^+(\varepsilon_i)(\varepsilon - \varepsilon_i) + \frac{\tilde{\mu}_{\max}}{2}(\varepsilon - \varepsilon_i)^2 \\ &= u_{\varepsilon_i}(\varepsilon) \end{aligned} \quad (113)$$

for $\varepsilon - \varepsilon_i \geq 0$. The first equality comes from claim 3, while the inequality comes from the fact that, according to claim 1 and the assumptions, $\dot{u}_{\varepsilon_{i-1}}(\varepsilon) \geq \dot{\tilde{\varphi}}^-(\varepsilon_i) \geq \dot{\tilde{\varphi}}^+(\varepsilon_i)$. However, we already know that $u_{\varepsilon_{i-1}}(\varepsilon) \geq \tilde{\varphi}(\varepsilon)$ for $\varepsilon \in [\varepsilon_{i-1}, \varepsilon_i]$ and $u_{\varepsilon_i}(\varepsilon) \geq \tilde{\varphi}(\varepsilon)$ for $\varepsilon \in [\varepsilon_i, \varepsilon_{i+1}]$. The claim follows.

■